OIPE JC26
APR 21 2003
PATENT & TRADEMARK OFFICE

AF/2700/8
#14
a/f
7/22/03

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

First Named Inventor : Venolia et al.	Appeal No. ---
Appln. No. : 09/421,710	
Filed : October 20, 1999	Group Art Unit: 2641
For : METHOD AND APPARATUS FOR DISPLAYING SPEECH RECOGNITION PROGRESS	Examiner: Armstrong
Docket No.: M61.12-0144	

RECEIVED
APR 24 2003
Technology Center 2600

TRANSMITTAL OF APPEAL BRIEF
(PATENT APPLICATION - 37 C.F.R. § 192)

Commissioner for Patents
Washington, D.C. 20231

I HEREBY CERTIFY THAT THIS PAPER IS BEING SENT BY U.S. MAIL, FIRST CLASS, TO THE ASSISTANT COMMISSIONER FOR PATENTS, WASHINGTON, D.C. 20231, THIS
17th DAY OF April, 2003
PATENT ATTORNEY

Sir:

Transmitted herewith in triplicate is the Appeal Brief in this application with respect to the Notice of Appeal filed on March 3, 2003.

FEE STATUS

[---] Small entity status under 37 C.F.R. §§ 1.9 and 1.27 is established by a verified statement.

FEE FOR FILING APPEAL BRIEF

Pursuant to 37 C.F.R. 1.17(c) the fee for filing the Appeal Brief is $320.00.

The Commissioner is authorized to charge any additional fees associated with this paper or credit any overpayment to Deposit Account No. 23-1123. A duplicate copy of this communication is enclosed.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

By: _____
Theodore M. Magee, Reg. No. 39,758
Suite 1600 - International Centre
900 Second Avenue South
Minneapolis, Minnesota 55402-3319
Phone: (612) 334-3222 Fax: (612) 334-3312

BEST AVAILABLE COPY

TMM:ajm:

OIPE
APR 21 2003
PATENT & TRADEMARK OFFICE

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

First Named Inventor : Venolia et al.	Appeal No. ---
Appln. No. : 09/421,710	
Filed : October 20, 1999	Group Art Unit: 2641
For : METHOD AND APPARATUS FOR DISPLAYING SPEECH RECOGNITION PROGRESS	Examiner: A Armstrong
Docket No.: M61.12-0144	

RECEIVED
APR 24 2003
Technology Center 2600

TRANSMITTAL OF APPEAL BRIEF
(PATENT APPLICATION - 37 C.F.R. § 192)

Commissioner for Patents
Washington, D.C. 20231

I HEREBY CERTIFY THAT THIS PAPER IS BEING SENT BY U.S. MAIL, FIRST CLASS, TO THE ASSISTANT COMMISSIONER FOR PATENTS, WASHINGTON, D.C. 20231, THIS 17th DAY OF April, 2003.

PATENT ATTORNEY

Sir:

Transmitted herewith in triplicate is the Appeal Brief in this application with respect to the Notice of Appeal filed on March 3, 2003.

FEE STATUS

[---] Small entity status under 37 C.F.R. §§ 1.9 and 1.27 is established by a verified statement.

FEE FOR FILING APPEAL BRIEF

Pursuant to 37 C.F.R. 1.17(c) the fee for filing the Appeal Brief is $320.00.

The Commissioner is authorized to charge any additional fees associated with this paper or credit any overpayment to Deposit Account No. 23-1123. A duplicate copy of this communication is enclosed.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

By: _____
Theodore M. Magee, Reg. No. 39,758
Suite 1600 - International Centre
900 Second Avenue South
Minneapolis, Minnesota 55402-3319
Phone: (612) 334-3222 Fax: (612) 334-3312

TMM:ajm:

#14

RECEIVED
APR 24 2003
Technology Center 2600

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

First Named Inventor	: Venolia et al.	Appeal No. ---
Appln. No.	: 09/421,710	
Filed	: October 20, 1999	Group Art Unit: 2641
For	: METHOD AND APPARATUS FOR DISPLAYING SPEECH RECOGNITION PROGRESS	Examiner: A. Armstrong
Docket No.:	M61.12-0144	

BRIEF FOR APPELLANTS

04/22/2003 YPOLITE1 00000136 09421710
01 FC:1402 320.00 OP

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

First Named Inventor	: Venolia et al.	Appeal No. ---
Appln. No.	: 09/421,710	
Filed	: October 20, 1999	Group Art Unit: 2641
For	: METHOD AND APPARATUS FOR DISPLAYING SPEECH RECOGNITION PROGRESS	Examiner: A. Armstrong
Docket No.:	M61.12-0144	

BRIEF FOR APPELLANT

Commissioner for Patents
Washington, D.C. 20231

I HEREBY CERTIFY THAT THIS PAPER IS BEING SENT BY U.S. MAIL, FIRST CLASS, TO THE ASSISTANT COMMISSIONER OF PATENTS, WASHINGTON, D.C. 20231, THIS 17th DAY OF April, 2003.

PATENT ATTORNEY

Sir:

This is an appeal of the final rejection of claims 1-33 that was reported in the Office Action of December 3, 2002.

REAL PARTY IN INTEREST

Microsoft Corporation, a corporation organized under the laws of the state of Washington, and having offices at One Microsoft Way, Redmond, Washington 98052, has acquired the entire right, title and interest in and to the invention, the application, and any and all patents to be obtained therefor, as assignee in an Assignment recorded on Reel 010578, frame 0070.

RELATED APPEALS AND INTERFERENCES

Appellant knows of no appeals or interferences that would directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

STATUS OF THE CLAIMS

Claims 1-33 are pending, rejected and appealed.

STATUS OF AMENDMENTS

No amendments were filed subsequent to the final rejection.

SUMMARY OF INVENTION

The invention provides a method and computer program for generating images of meters related to speech recognition on a display device 85.

Under one embodiment of the method, an insertion marker 214 is displayed on display 85 and a progress meter 244, 376 is displayed near the insertion marker. The progress meter quantitatively indicates the amount of progress in decoding a speech input. (See page 17, lines 17-20). A method of generating a progress meter is shown in Fig. 12 and discussed on pages 33-34.

Under other embodiments of the present invention, a quantitative progress meter is combined with a volume meter as shown in Figs. 4G, 4H, 4I, 6E, 7D, 8C and 9D. By combining the progress meter and the volume meter, the user is able to track the volume of their speech and the progress of the speech recognition system without substantially moving their eyes.

Under one such embodiment, the volume meter is constructed according to a method shown in Fig. 11. In Fig. 11, an input analog value is converted into a digital value that is then transformed (step 459 of Fig. 11). The transformed value is stored in a buffer at step 460. The value of the transformed value is compared to a transform maximum to determine a percentage value at step 466. This percentage value is then used to form the volume meter.

In some embodiments, each speech sample is associated

with a separate token, such as tokens 220, which consist of a set of subordinate blocks such as blocks 228 and 230 of Fig. 4D. Tokens also include a background rectangle 232. In some embodiments, the progress meter is shown by converting the color of the background rectangle of each volume token.

The size of the blocks in the tokens is determined at step 468. Up to three different blocks can be constructed for a single speech value with each block having a maximum size. The determination of the subordinate block sizes is discussed on page 29, line 12 to page 31, line 5, and involves calculating a height ratio based on the percentage of the transformed value to the transform maximum and comparing the height ratio to maximum height ratios for each block.

ISSUES

Does the combination of VanBuskirk et al. (U.S. Patent No. 6,075,534), Tannenbaum (U.S. Patent No. 6,233,560) and Rozak et al. (U.S. Patent No. 5,864,815) render claims 1-3, 13-14, 17-21 and 29 obvious under 35 U.S.C. § 103(a)?

Does the combination of VanBuskirk et al., Tannenbaum, Rozak et al. and French-St. George et al. (U.S. Patent No. 6,018,711) render claims 4-16, 20, 22-28 and 30-33 obvious under 35 U.S.C. § 103(a)?

GROUPING OF CLAIMS

Appellants group the claims on appeal as follows:

Group 1 - Claims 1, 2, 17 and 18

Group 2 - Claim 3

Group 3 - Claim 4

Group 4 - Claim 5

Group 5 - Claim 6

Group 6 - Claim 7

Group 7 - Claim 8

Group 8 - Claim 9
Group 9 - Claim 10
Group 10 - Claim 11
Group 11 - Claim 12
Group 12 - Claim 13
Group 13 - Claim 14
Group 14 - Claim 15
Group 15 - Claim 16
Group 16 - Claims 19
Group 17 - Claim 20
Group 18 - Claims 21 and 29
Group 19 - Claims 22 and 23
Group 20 - Claim 24
Group 21 - Claim 25
Group 22 - Claim 26
Group 23 - Claim 27
Group 24 - Claim 28
Group 25 - Claims 30 and 31
Group 26 - Claim 32
Group 27 - Claim 33

ARGUMENT

Group 1 - Claims 1, 2, 17 and 18

Independent claims 1 and 17 and dependent claims 2 and 18 were rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk et al. (U.S. Patent No. 6,075,537, hereinafter VanBuskirk) in view of Tannenbaum (U.S. Patent No. 6,233,560) and Rozak et al. (U.S. Patent No. 5,864,815, hereinafter Rozak).

Independent claim 1 provides a method of displaying images on a display device. The method includes displaying an insertion marker at an insertion area on a display and displaying a progress meter near the insertion area based on the location of

the insertion marker. The progress meter quantitatively indicates the amount of progress in decoding a speech input.

Independent claim 17 provides a computer program having a least one insertion point marker that indicates a location on the display where a user desires to provide input. The computer program also includes a speech recognition routine and a meter generation routine that displays a progress meter near an insertion point based on the insertion point marker. The progress meter is quantitatively indicative of the amount of speech that has been decoded by the speech recognition routine.

The combination of VanBuskirk, Tannenbaum and Rozak does not show or suggest the invention of claim 1 or claim 17 because none of these references show a progress meter that quantitatively indicates the amount of progress in decoding a speech input. In addition, none of these references show or suggest the ability to place such a progress meter at an insertion area.

VanBuskirk describes a volume tracking window for a speech recognition system. Under VanBuskirk, the detected volume of a speech signal is represented in a window by changing the color of the entire window or by moving a colored bar horizontally to show the current volume. The volume tracking window in VanBuskirk may be a "floating window", however, VanBuskirk does not show or suggest that the volume tracking window should be placed near an insertion marker. In addition, VanBuskirk does not show or suggest a progress meter that shows the amount of progress in decoding an input speech signal.

Tannenbaum discloses a speech recognition interface in which fully recognized phrases or commands are displayed in a box. Under Tannenbaum, the recognized phrase or command is used to determine where to position the box on the display. Tannenbaum does not disclose displaying a progress meter that indicates the

amount of progress in recognizing a speech signal. In addition, because Tannenbaum requires the recognition of the command before it can determine where to place the recognition result, its teachings could not be used to position a recognition progress meter since the commands would not be known during the period of time when the progress meter needs to be displayed. Thus, under Tannenbaum, it would not be possible to place a progress meter near an insertion area since the command is unknown until recognition is complete.

Rozak discloses a graphical notification that indicates the processing state of a speech recognizer. The graphical notification provides an icon showing whether the speech recognizer is currently processing audio inputs. However, Rozak does not teach or suggest a progress meter that quantitatively indicates the amount of progress in decoding a speech input. Rozak also does not suggest placing a progress meter near an insertion area.

In the Final Office Action, the Examiner asserted that even though no single reference showed a quantitative progress meter, it would have been obvious to combine the quantitative aspect of the volume meter in VanBuskirk with the non-quantitative status indicator of Rozak to produce a quantitative meter that is indicative of the amount of speech recognition that has been performed. Appellants dispute this assertion.

First, there is no suggestion in Rozak or VanBuskirk for taking the quantitative attribute of the volume meter provided by VanBuskirk and combining it with the state information shown by Rozak. As stated by the Federal Circuit in In re Lee, 277 F.3d 1338, 61 U.S.P.Q.2d 1430, 1432 (Fed. Cir. 2002),"particular findings must be made as to the reason the skilled artisan, with no knowledge of the claimed invention, would have selected these components for combination in the manner claimed" (citing In re

Rouffet, 149 F.3d 1350, 1359, 47 U.S.P.Q.2d 1453, 1459 (Fed. Cir. 1998)). However, in VanBuskirk and Rozak there are no such suggestions. VanBuskirk does not indicate that the quantitative aspect of its volume meter should be applied to other types of display meters and Rozak does not indicate that there are any problems with its non-quantitative meter that would require it to be changed.

In fact, the selective combination of these aspects of VanBuskirk and Rozak was only made by the Examiner after reading the blueprint provided by the present application. Without the description of the present invention provided by this application, neither the Examiner nor those skilled in the art would have thought to combine the quantitative aspect of the volume meter of VanBuskirk with the basic status display of Rozak. The Examiner is clearly using hindsight in forcing the combination of these two references to produce the invention described in the present application. It is this type of hindsight that the Federal Circuit has attempted to limit by requiring that the Examiner find some suggestion for selecting particular components for combination.

In addition, Appellants note that this selective combination does not teach those skilled in the art how to practice the present invention. In particular, neither VanBuskirk nor Rozak teach a manner in which a progress meter that quantitatively indicates the amount of progress in decoding a speech input could be displayed. In VanBuskirk, the quantitative amount of the volume is easily displayed by selecting a maximum value to be expected, and displaying a ribbon based on the ratio between the current volume and the maximum volume. However, neither VanBuskirk nor Rozak show a way to set a maximum value to indicate the progress of decoding a speech input. In particular, neither reference shows how to determine how large to make the progress meter at any point in time or how to change the progress

meter over time so that it quantitatively shows the progress in decoding a speech input. Without such teachings, it is not clear how such a progress meter could be formed from VanBuskirk and Rozak.

Since there is no suggestion to combine the selected elements of VanBuskirk and Rozak, and since the references do not teach how to generate a quantitative progress meter, it would not be obvious to combine the selected elements in VanBuskirk and Rozak to form the invention of claims 1, 2, 17 and 18.

In addition, the combination of the cited references do not show or suggest the ability to place a progress meter at an insertion area designated by an insertion marker. In the Final Office Action, it was asserted that Tannenbaum teaches this aspect of the invention because it teaches that recognized commands should be displayed at a location functionally related to the analyzed contents and context of the voice input. However, the teachings of Tannenbaum could not be applied to a progress meter. In particular, Tannenbaum's system requires that the input be analyzed to determine which command has been spoken. See column 2, lines 40-47 and column 7, lines 7-24. Note in particular that the system does not know if the user has issued a command to open a file or move a cursor until it has decoded the word spoken by the user. Without that information, Tannenbaum does not know whether to place the feedback near the tool bar or near the cursor. Thus, Tannenbaum must wait for the recognition to be performed before placing its feedback box on the screen.

However, waiting to place the feedback box until after the recognition is complete defeats the purpose of displaying a progress meter indicating the progress of recognition. In particular, since Tannenbaum must wait for the full recognition to finish, it is impossible for Tannenbaum's system to be used to place a progress meter that must be displayed before the

recognition is completed.

In light of the fact that none of the cited references provide a system that allows a progress meter to be positioned at an insertion marker and because there is no suggestion in the cited references for displaying a quantitative progress meter or suggestions of how a progress meter could be constructed, the invention of claims 1, 2, 17 and 18 is patentable over VanBuskirk, Rozak and Tannenbaum.

Group 2 - Claim 3

Dependent claim 3 depends from claim 1. It does not stand or fall with claim 1 because it includes an additional limitation that is not shown or suggested in the combination of VanBuskirk, Rozak and Tannenbaum.

In claim 3, a user's speech input is converted into an analog speech signal and that signal is converted into at least one digital speech value. The at least one digital speech value is then transformed into coordinates for at least one shape on the display which is positioned near the progress meter.

Claim 3 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, and Tannenbaum.

The combination of VanBuskirk, Rozak and Tannenbaum does not show or suggest a display that includes both a progress meter and a shape that is formed by transforming a digital speech value into coordinates.

In the Final Office Action, it was asserted that VanBuskirk teaches that a multi-function graphical user interface should be shown in as small a space as possible and that it would therefore be obvious to display a volume meter near a progress meter. Applicants dispute this assertion.

First, Appellants note that none of these references show or suggest a progress meter that quantitatively shows the

amount of progress in decoding a speech signal. Thus, those skilled in the art would first have to create such a progress meter from these references without even a suggestion in any of the references that such a progress meter should be constructed.

After the progress meter is constructed, those skilled in the art still would not have produced the invention of claim 3. To produce the invention of claim 3, those skilled in the art would have to take the further step of combining the progress meter with another shape formed by transforming a digital speech value. However, there is no suggestion in the art for combining a progress meter with such a shape in the cited art.

Although VanBuskirk does teach that different types of information can be conveyed in the same graphical interface, it does not show or suggest that a shape formed by transforming a digital speech value should be combined with a progress meter.

Since none of the cited references show or suggest a progress meter, the combination of references cannot show or suggest the combination of a progress meter and a shape generated from a digital speech value. As such, claim 3 is patentable over VanBuskirk, Rozak and Tannenbaum.

Group 3 - Claim 4

Claim 4 depends from claim 3. Claim 4 does not stand or fall with claim 3 because it includes an additional limitation that is not found in the combination of VanBuskirk, Rozak and Tannenbaum.

In claim 4, the step of transforming at lest one digital speech value into coordinates includes applying a mathematical function to the at least one digital speech value to produce a transform value and using the transform value to identify coordinates for the shape on the display.

Claim 4 was rejected under 35 U.S.C. § 103(a) as being

unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George et al. (U.S. Patent No. 6,018,711, hereinafter French-St. George).

French-St. George discloses an animation that indicates the amount of time the user has left in which to provide speech input to a speech recognizer. French-St. George does not suggest that this animation should be placed near an insertion point and does not show or suggest a progress meter that indicates the amount of progress in decoding an input speech segment. In addition, French-St. George does not suggest identifying coordinates for at least one shape on the display by applying a mathematical function to at least one digital speech value to produce a transform value.

In fact, none of the references in the combination of VanBuskirk, Rozak, Tannenbaum, and French-St. George show or suggest applying a mathematical function to at least one digital speech value to produce a transform value and then using that transform value to identify coordinates for at least one shape on a display. In rejecting claim 4, the Examiner never asserted that any of the references show a mathematical function applied to at least one digital speech value to produce a transform value. As such, the invention of claim 4 is not shown or suggested in the combination of the references cited by the Examiner and claim 4 is therefore patentable over VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 4 - Claim 5

Claim 5 depends from 4. Claim 5 does not stand or fall with claim 4 because it includes a further limitation wherein applying a mathematical function comprises taking the logarithm value of at least one digital value. This additional limitation is not shown or suggested in the combination of VanBuskirk, Rozak,

Tannenbaum and French-St. George.

Claim 5 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George. However, none of the cited references show or suggest taking a logarithm of a digital speech value. Furthermore, the Examiner has not asserted that any particular section of any of these references shows the taking of a logarithm of a digital speech value.

Since none of the references show or suggest taking the logarithm of a speech value, claim 5 is patentable over the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 5 - Claim 6

Claim 6 depends from 4 and includes a further limitation wherein the mathematical function comprises taking the square root of at least one digital value. Claim 6 does not stand or fall with claim 4 because the additional limitation of taking a square root of a digital speech value is not shown in the cited references.

Claim 6 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George.

None of the cited references show or suggest taking the square root of at least one digital speech value to produce a transform value that is then used to identify the coordinates of at least one shape on the display. In addition, no section of the references has been cited as showing the application of a square root function to a digital speech value.

As such, claim 6 is patentable over the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 6 - Claim 7

Claim 7 depends from claim 4 but stands apart from claim 4 because it includes a further limitation wherein the shape formed from the transformed value is a base rectangle and determining the coordinates of the base rectangle includes a number of steps that are not shown in the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George.

In claim 7, coordinates of the base rectangle are determined by determining a base point for the base rectangle on the display, accessing a stored rectangle width, and accessing a maximum transform value. The transform value formed from the digital speech value is divided by the maximum transform value to produce a transform ratio. A height is calculated based in part on the transform ratio. The coordinates of the base rectangle are then calculated based on the base point, the stored rectangle width, and the calculated height.

Claim 7 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George. However, none of the cited references show or suggest the steps of claim 7. In particular, none of the cited references show or suggest accessing a maximum transform value, dividing a transform value produced from a digital speech value by the maximum transform value, or using the resulting ratio to calculate coordinates for a base rectangle.

Since none of these steps are shown in any of the references, the invention of claim 7 is patentable over VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 7 - Claim 8

Claim 8 depends from claim 7 and includes a further limitation where determining a calculated height includes determining if the transform ratio is greater than a maximum

height ratio for the base rectangle and if it is, performing a further step of multiplying the maximum height ratio for the base rectangle by the full meter height to produce the calculated height. Since this limitation is not shown in the combination of VanBuskirk, Rozak, Tannenbaum, and French-St. George, claim 8 does not stand or fall with claim 7.

Claim 8 was rejected under 35 U.S.C. § 103(a) as being unpatentable over the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George. However, none of the cited references show or suggest the steps of claim 8. In particular, none of the references show a step of comparing a transform ratio to a maximum height ratio for a base rectangle to determine which is larger. In addition, none of the references show or suggest performing a further step of multiplying the maximum height ratio for the base rectangle by a full meter height if the transform ratio is greater than the maximum height ratio.

Since none of the cited references show these steps, the combination of these references does not show or suggest these steps. As such, claim 8 is patentable over the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 8 - Claim 9

Claim 9 depends from claim 8 and was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French St. George. Claim 9 does not stand or fall with claim 8 because it includes an additional limitation to determine the coordinates of a second rectangle based on the coordinates of a base rectangle, which is not shown or suggested in the cited art.

In claim 9, the maximum height ratio for the base rectangle is subtracted from the transform ratio to produce an excess ratio. A second rectangle height is then determined based

in part on the excess ratio. The coordinates of a second rectangle are then calculated based on the coordinates of the base rectangle, a stored rectangle width, and the second rectangle height.

None of the cited references show a step of subtracting a maximum height ratio from a transform ratio or a step of using the resulting difference to determine the height of a second rectangle. As such, the combination of references does not show the invention of claim 9. Claim 9 is therefore patentable over VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 9 - Claim 10

Claim 10 depends from claim 9 and includes a further limitation wherein determining the second rectangle height involves comparing the excess ratio to a maximum height ratio for the second rectangle. Since this additional limitation is not shown in the references used to reject claim 10, claim 10 does not stand or fall with claim 9.

Claim 10 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George. However, no particular section of any of these references was cited by the Examiner to support the rejection of claim 10 and in fact, no section of these references shows a step of determining a rectangle height for a second rectangle. As such, claim 10 is patentable over VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 10 - Claim 11

Claim 11 depends from claim 10 but does not stand or fall with claim 10 because it includes an additional limitation to determining the coordinates of a third rectangle, which is not shown or suggested in the cited art.

Claim 11 was rejected under 35 U.S.C. § 103(a) as being obvious from VanBuskirk, Rozak, Tannenbaum and French-St. George.

Under claim 11, the maximum height ratio for the second rectangle is subtracted from the excess ratio to produce a remainder ratio. A third rectangle height is then determined by multiplying the remainder ratio by the full meter height. The coordinates of a third rectangle is then determined from the third rectangle height.

None of the references in the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George shows or suggests determining the coordinates of a third rectangle. In the Final Office Action, no portion of any of these references was cited as disclosing the calculation of a third rectangle. As such, the invention of claim 11 is patentable over the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 11 - Claim 12

Claim 12 depends from claim 7. Claim 12 does not stand or fall with claim 7 because it includes a further limitation of calculating the coordinates of a background rectangle that is not found in the cited references.

Claim 12 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George.

Claim 12 includes a limitation to calculating the coordinates of a background rectangle where the background rectangle appears somewhere between the base rectangle and a point at a full meter height above a bottom edge of the base rectangle. None of the cited references disclose calculating the coordinates of such a background rectangle. In addition, the Final Office Action did not cite any portion of these references as showing such a background rectangle. As such, the invention of claim 12 is

patentable over the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 12 - Claim 13

Claim 13 depends from independent claim 1. Claim 13 does not stand or fall with claim 1 because it includes a further limitation that is not shown or suggested in the cited references.

Claim 13 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak and Tannenbaum and as being unpatentable over the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George.

In claim 13, the progress meter of claim 1 is displayed by dividing the speech input into frames, and decoding at least one of the frames of the speech input into a sub-word unit. A frame number for the last frame to be decoded is divided by the total number of frames to produce a decode ratio. The progress meter is then displayed based on this decode ratio.

None of the cited references show or suggest a step of dividing the frame number of a last frame to be decoded by the total number of frames to produce a decode ratio. Further, none of the references show or suggest using a decode ratio to display a progress meter. As such, claim 13 is patentable over the combination of VanBuskirk, Rozak and Tannenbaum and the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 13 - Claim 14

Claim 14 depends from claim 13. Claim 14 does not stand or fall with claim 13 because it includes a further limitation that is not shown in the cited references.

Like claim 13, claim 14 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak and Tannenbaum

and as being unpatentable over the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George.

Claim 14 includes limitations to multiplying the decode ratio by a full meter width to determine a progress width and calculating the coordinates of a progress rectangle based on the progress width, a stored meter height and a base point on the display. None of the cited references show or suggest multiplying a decode ratio by a full meter width, nor do they show or suggest calculating the coordinates of a progress rectangle based on the result of such a multiplication. Since none of the references show any of these steps, their combination does not show or suggest the invention of claim 14, and claim 14 is therefore patentable over VanBuskirk, Rozak and Tannenbaum and the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 14 - Claim 15

Claim 15 depends from claim 12. It does not stand or fall with claim 12 because it includes further limitations that are not found in the cited references.

Claim 15 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George. However, no particular section of any of these references was cited as showing or suggesting any of the elements of claim 15.

In claim 15, the speech input is divided into frames and at least one frame of the speech is decoded into a sub-word unit. A frame number for the last frame to be decoded is divided by the total number of frames to produce a decode ratio. A progress meter is then displayed based on the decode ratio by changing the color of at least one background rectangle.

None of the cited references discuss dividing a frame number of the last frame to be decoded by the total number of

frames to produce a decode ratio. In addition, none of the references discuss using the decode ratio to change the color of at least one background rectangle and thereby display a progress meter.

Since none of the references discuss these steps, their combination does not show or suggest the steps of claim 15. As such, claim 15 is patentable over the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 15 - Claim 16

Claim 16 depends from claim 15 but does not stand or fall with claim 15 because it includes a further limitation that is not shown or suggested in the cited art.

Claim 16 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George.

In claim 16, the step of displaying a progress meter in claim 15 is defined as multiplying the decode ratio by a full meter width to produce a progress width and dividing the progress width by a rectangle width that is indicative of the width of each background rectangle. This produces a rectangle count. The color of a number of background rectangles is then changed where the number of background rectangles is equal to the rectangle count.

None of the cited references show or suggest the steps of dividing a progress width by a rectangle width to get a rectangle count or of changing the color of a number of background rectangles equal to the rectangle count. As such, the combination of references does not show these steps.

Since none of the references show the steps of claims 16, the invention of claim 16 is patentable over VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 16 - Claim 19

Claim 19 depends from claim 17 but does not stand or fall with claim 17 because it includes an additional limitation that is not found in the cited art.

In particular, claim 19 includes a limitation where a meter generation routine further comprises a transform routine that transforms a digital value into a set of coordinates for a shape on the display. The digital value is indicative of the magnitude of a portion of the speech signal. Thus, the meter generation routine of claim 19 is able to generate a progress meter that quantitatively indicates the progress of decoding while at the same time providing a shape on the display that indicates the magnitude of the speech signal.

Claim 19 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak and Tannenbaum.

In the Final Office Action, it was noted that none of these references specifically teach displaying a volume meter close to a progress meter. However, it was asserted that since VanBuskirk teaches that a multi-function graphical user interface should supply information in the smallest possible space, it would be obvious to display a volume meter close to a decoding progress meter. Appellants dispute this assertion.

Appellants note that none of these references show or suggest a progress meter that quantitatively shows the amount of progress in decoding a speech signal. Thus, those skilled in the art would first have to create such a progress meter without any suggestion from these references to do so.

After the progress meter is constructed, those skilled in the art still would not have produced the invention of claim 19. To produce the invention of this claim, those skilled in the art would have to take a further step of combining the progress meter with a shape that indicates the magnitude of a speech value.

However, there is no suggestion in the art for making such a combination.

Although VanBuskirk does teach that different types of information can be conveyed in the same graphical interface, it does not show or suggest that a volume meter should be combined with a progress meter. In addition, none of the references disclose how a progress meter could be displayed near a volume meter without confusing the user as to what the meters are conveying. Thus, there is no teaching or suggestion in any of the cited art for how to place a volume meter near a progress meter.

In light of the fact that none of the references show a progress meter that shows the amount of progress in decoding a speech input and that none of the references show or suggest combing such a progress meter with a shape representing the magnitude of a speech value, the invention of claim 19 is patentable over the combination of VanBuskirk, Rozak and Tannenbaum.

Group 17 - Claim 20

Claim 20 depends from claim 17 but does not stand or fall with claim 17 because it was rejected based on a different combination of references from claim 17. In particular, claim 20 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George.

Claim 20 includes a further limitation wherein the speech recognition routine decodes the speech signal into a set of subwords.

As discussed above for claim 17, the combination of VanBuskirk, Rozak, and Tannenbaum does not show a quantitative progress meter and does not suggest how a quantitative progress meter can be constructed.

French-St. George also fails to show a quantitative

progress meter that shows the amount of progress in decoding a speech signal. French-St. George only shows an animation related to the amount of time left for a user to provide speech to a speech recognition system. It does not discuss the progress of decoding a speech signal.

Since French-St. George does not disclose a quantitative progress meter that indicates the amount of progress in decoding a speech signal and none of VanBuskirk, Rozak, and Tannenbaum show such a meter, the invention of claim 20 is patentable over the combination of VanBuskirk, Rozak, Tannenbaum, and French-St. George.

Group 18 - Claims 21 and 29

Independent claims 21 and 29 stand apart from independent claims 1 and 17 because they include a combination of limitations that are not found in independent claims 1 and 17.

Independent claim 21 is directed to a method in a computer system that displays a volume meter that is indicative of the magnitude of at least a portion of a speech input. The volume meter is displayed near a progress meter, where the progress meter quantitatively indicates the amount of progress in decoding the speech signal.

Independent claim 29 provides a computer program designed to operate in a computer system having a display. The program includes a volume meter portion that is capable of displaying a volume meter, a speech recognition portion that is capable of converting human speech into a set of sub-words, and a progress meter portion that is capable of generating a progress meter that is quantitatively indicative of the amount of progress in converting the human speech signal into sub-words.

Claims 21 and 29 were rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, and Tannenbaum.

In the Final Office Action, the Examiner stated that none of VanBuskirk, Tannenbaum or Rozak specifically teaches displaying a volume meter close to a progress meter. However, the Examiner asserted that VanBuskirk teaches that multiple functions should be placed in a single graphical user interface so that the interface is as small as possible and that because of this, it would be obvious to combine a volume meter with a progress meter. Appellants dispute this assertion.

Appellants note that none of these references show or suggest a progress meter that quantitatively shows the amount of progress in decoding a speech signal. Thus, those skilled in the art would first have to create such a progress meter without even a suggestion in the references that such a progress meter should be constructed and without any teachings as to how to construct such a meter.

After the progress meter is constructed, those skilled in the art still would not have produced the invention of claims 21 and 29. To produce the invention of those claims, those skilled in the art would have to take the further step of combining the progress meter with a volume meter so that the two meters are displayed close together. Once again, there is no suggestion in the art for combining a volume meter with a progress meter.

Although VanBuskirk does teach that different types of information can be conveyed in the same graphical interface, it does not show or suggest that a volume meter should be combined with a progress meter. In addition, none of the references disclose how a progress meter could be displayed near a volume meter without confusing the user as to what the meters are conveying. Thus, there is no teaching or suggestion in any of the cited art for how to place a volume meter near a progress meter.

In light of the fact that none of the references show or suggest a quantitative progress meter and therefore are

incapable of showing or suggesting that such a progress meter should be placed near a volume meter, the combination of references does not show or suggest the invention of claims 21 and 29. As such, claims 21 and 29 are patentable over VanBuskirk, Rozak and Tannenbaum.

Group 19 - Claims 22 and 23

Claim 22 depends from claim 21 and claim 23 depends from claim 22. Claims 22 and 23 do not stand or fall with claim 21 because they include a further limitation that is not shown in the cited art and because they were rejected based on a different combination of art.

Claim 22 includes a further limitation wherein displaying a volume meter includes storing digital values representing the magnitudes of different portions of the speech signal. A separate token is then displayed for each separate digital value.

Claims 22 and 23 were rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George. No portions of any of these references were cited as showing a step of displaying a separate token for separate digital values representing the magnitude of a speech signal.

In fact, none of the references show or suggest displaying a separate token for separate digital values of a speech signal. Since none of the cited references show displaying a separate token, claims 22 and 23 are patentable over VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 20 - Claim 24

Claim 24 depends from claim 23. It does not stand or fall with claim 23 because it includes an additional limitation

that is not shown in the cited references.

In claim 24, a digital value representing the magnitude of a portion of a speech signal is transformed into a transform value. This transform value is then divided by a maximum meter value to produce a meter ratio. The height of a meter portion of a token is then determined using the meter ratio and the full meter height.

Claim 24 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George. However, no section of any of these references was cited as showing a step of transforming a digital value into a transform value and then dividing the transform value by a maximum meter value.

In fact, none of the references show or suggest these steps. As such, the invention of claim 24 is patentable over VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 21 - Claim 25

Claim 25 depends from claim 24 and includes a further limitation wherein determining the height comprises multiplying the meter ratio by the full meter height. Claim 25 does not stand or fall with claim 24 because this additional limitation is not found in the cited references.

Claim 25 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George. However, none of these references show or suggest multiplying a meter ratio by a full meter height to determine the height of a meter portion of a token. As such, the invention of claim 25 is patentable over VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 22 - Claim 26

Claim 26 depends from claim 24 and includes a further limitation wherein determining the height of a meter portion of a token includes performing a series of steps that are not shown in the cited references. As such, claim 26 does not stand or fall with claim 24.

Claim 26 includes steps of determining if the meter ratio is greater than a base ratio and if the meter ratio is greater, multiplying the base ratio by the full meter height to determine the height of a base block while subtracting the base ratio from the meter ratio to produce an excess ratio. This excess ratio is then used with the full meter height to determine the height of a second block of the meter portion.

Claim 26 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George. However, no particular portion of any of these references was cited to support this rejection.

None of the cited references mention any of the steps of comparing a meter ratio to a base ratio, multiplying a base ratio by a full meter height, subtracting a base ratio from a meter ratio to produce an excess ratio, or using an excess ratio and the full meter height to determine the height of a second block. As such, the combination of references does not show these steps. In light of this, claim 26 is patentable over the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 23 - Claim 27

Claim 27 depends from claim 26 but does not stand or fall with claim 26 because it includes further steps that are not shown in the cited references.

In particular, claim 27 includes steps of comparing the

excess ratio to an intermediate ratio to determine if it exceeds the intermediate ratio and if it does exceed the intermediate ratio multiplying the intermediate ratio by the full meter height to produce the height of the second block. The intermediate ratio is then subtracted from the excess ratio to produce a remainder ratio. This is multiplied by the full meter height to produce the height for a top block of the meter portion.

Claim 27 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George. However, none of these references show or suggest any of the steps found in claim 27. As such, claim 27 is patentable over the cited references.

Group 24 - Claim 28

Claim 28 depends from claim 24 but does not stand or fall with claim 24 because it includes further limitations that are not found in the cited references.

In particular, claim 28 includes steps of dividing a number for a last decoded frame by a total number of frames to produce a progress ratio and multiplying the progress ratio by a full meter width to produce a progress width. The progress width is then divided by a token width to produce an affected number of tokens. For each of the affected number of tokens, the color of at least a portion of the token is changed so that it is different from the color of other tokens. Thus, claim 28 displays the progress meter by modifying the tokens used for the volume meter.

Claim 28 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George. No portions of any of these references were cited as showing any of the steps of claim 28.

In fact, none of these references show any of these steps. As such, the combination of these references does not show

or suggest the steps of claim 28. Therefore, claim 28 is patentable over VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 25 - Claims 30 and 31

Claim 30 depends from claim 29 and claim 31 depends from claim 30. Claims 30 and 31 do not stand or fall with claim 29 because they were rejected using a different combination of references than those used to reject claim 29.

Claims 30 and 31 were rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George. Since claims 30 and 31 depend from claim 29, they include the limitations to displaying a progress meter near a volume meter that are found in claim 29.

As discussed above for claim 29, the combination of VanBuskirk, Rozak, and Tannenbaum does not show or suggest a quantitative progress meter in combination with a volume meter. Similarly, French-St.George does not show or suggest a quantitative progress meter that shows the amount of progress in decoding a speech signal and does not show or suggest combining such a progress meter with a volume meter. As such, the combination of VanBuskirk, Rozak, Tannenbaum and French-St.George does not show the invention of claim 29 or claims 30 and 31, which depend from claim 29. Therefore, claims 30 and 31 are patentable over VanBuskirk, Rozak, Tannenbaum and French-St.George.

Group 26 - Claim 32

Claim 32 depends from claim 31 but does not stand or fall with claim 31 because it includes a further limitation that is not shown or suggested in the cited art.

In particular, claim 32 includes components for taking a frame number representing the last frame of speech that was

decoded by a speech recognition system and dividing it by a total frame number representing the total number of frames found in the speech signal. It also includes a component for determining a maximum dimension for the progress meter and code for multiplying the progress ratio by the maximum dimension to produce a progress dimension.

Claim 32 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George. In the Final Office Action, no particular sections of any of these references were cited as showing the functions described in claim 32.

In fact, none of the cited references show or suggest any of the functions performed in claim 32. As such, the combination of these references does not show or suggest the invention of claim 32. Claim 32 is therefore patentable over VanBuskirk, Rozak, Tannenbaum and French-St. George.

Group 27 - Claim 33

Claim 33 depends from claim 32 but does not stand or fall with claim 32 because it includes a further limitation that is not found in the cited references. In particular, claim 33 includes a limitation wherein volume token program code generates a volume token based in part on the progress dimension.

Claim 33 was rejected under 35 U.S.C. § 103(a) as being unpatentable over VanBuskirk, Rozak, Tannenbaum and French-St. George. No particular portions of those references were cited to support the rejection of claim 33. In particular, no portions were cited as showing the generation of a volume token based in part on a progress dimension.

The invention of claim 33 is not shown or suggested in any of the cited references. In particular, none of the references show or suggest that a volume token should be based in part on a

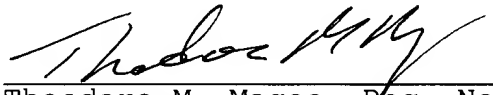
progress dimension. As such, the invention of claim 33 is patentable over the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George.

CONCLUSION

In conclusion, claims 1-33 are not shown or suggested in the combination of VanBuskirk, Rozak, Tannenbaum and French-St. George. Appellants therefore request reversal of the Examiner's rejection of claims 1-33.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

By: 

Theodore M. Magee, Reg. No. 39,758
Suite 1600 - International Centre
900 Second Avenue South
Minneapolis, Minnesota 55402-3319
Phone:(612) 334-3222 Fax:(612) 334-3312

TMM:ajm

Appendix A

Claims On Appeal:

1. A method in a computer system for generating images on a display device, the method comprising:

displaying an insertion marker at an insertion area on a display, the insertion area representing the location at which the user desires to provide input; and

displaying a progress meter near the insertion area based on the location of the insertion marker, the progress meter quantitatively indicative of the amount of progress in decoding a speech input.

2. The method of claim 1 further comprising:

receiving input from the user indicating that a microphone is to be activated;

activating the microphone; and

displaying an indication that the microphone is active near the progress meter.

3. The method of claim 1 further comprising:

converting a user's speech input into an analog speech signal;

converting the analog speech signal into at least one digital speech value; and

transforming the at least one digital speech value into coordinates for at least one shape on the display positioned near the progress meter.

4. The method of claim 3 wherein transforming the at least one digital speech value into coordinates comprises:

applying a mathematical function to the at least one digital speech value to produce a transform value, the range between the lowest possible transform value and the highest possible transform value being less than the range between the lowest possible digital speech value and the highest possible speech value; and

using the transform value to identify coordinates for the at least one shape on the display.

5. The method of claim 4 wherein applying a mathematical function comprises taking the logarithm of at least one digital value.

6. The method of claim 4 wherein applying a mathematical function comprises taking the square-root of at least one digital value.

7. The method of claim 4 wherein the shape is a base rectangle and wherein using the transform value to determine the coordinates of the base rectangle comprises:

determining a base point for the base rectangle on the display;

accessing a stored rectangle width;

accessing a maximum transform value;

dividing the transform value by the maximum transform value to produce a transform ratio;

determining a calculated height based in part on the transform ratio; and

calculating the coordinates of the base rectangle based on the base point, the stored rectangle width and the calculated height.

8. The method of claim 7 wherein determining a calculated height comprises determining if the transform ratio is greater than a maximum height ratio for the base rectangle and if it is, performing a further step of multiplying the maximum height ratio for the base rectangle by a full meter height to produce the calculated height.

9. The method of claim 8 further comprising steps of:

subtracting the maximum height ratio for the base rectangle from the transform ratio to produce an excess ratio;

determining a second rectangle height based in part on the excess ratio; and

calculating the coordinates of a second rectangle based on the coordinates of the base rectangle, the stored rectangle width and the second rectangle height, the coordinates of the base rectangle and the second rectangle such that the second rectangle appears connected to a top edge of the base rectangle on the display.

10. The method of claim 9 wherein determining a second rectangle height comprises determining if the excess ratio is greater than a maximum height ratio for the second rectangle and if it is, performing a further step of multiplying the maximum height ratio for the second rectangle by the full meter height to produce the second rectangle height.

11. The method of claim 10 further comprising steps of:

subtracting the maximum height ratio for the second rectangle from the excess ratio to produce a remainder ratio;

determining a third rectangle height by multiplying the remainder ratio by the full meter height; and

calculating the coordinates of a third rectangle based on the coordinates of the second rectangle, the stored rectangle width and the third rectangle height, the coordinates of the second rectangle and the third rectangle such that the third rectangle appears connected to a top edge of the second rectangle on the display.

12. The method of claim 7 further comprising calculating the coordinates of a background rectangle, the background rectangle appearing somewhere between the base rectangle and a point at a full meter height above a bottom edge of the base rectangle.

13. The method of claim 1 wherein displaying a progress meter further comprises:

dividing the speech input into frames;

decoding at least one of the frames of speech into a sub-word unit;

dividing a frame number of the last frame to be decoded by the total number of frames to produce a decode ratio; and

displaying the progress meter based on the decode ratio.

14. The method of claim 13 wherein displaying the progress meter further comprises:

multiplying the decode ratio by a full meter width to determine a progress width; and

calculating the coordinates of a progress rectangle based on the progress width, a stored meter height and a base point on the display.

15. The method of claim 12 further comprising:

dividing the speech input into frames;

decoding at least one of the frames of speech into a sub-word unit;

dividing a frame number of the last frame to be decoded by the total number of frames to produce a decode ratio; and

displaying the progress meter based on the decode ratio by changing the color of at least one background rectangle.

16. The method of claim 15 wherein displaying the progress meter comprises:

multiplying the decode ratio by a full meter width to produce a progress width;

dividing the progress width by a rectangle width that is indicative of the width of each background rectangle, the division producing a rectangle count; and

changing the color of a number of background rectangles, the number of background rectangles being equal to the rectangle count.

17. A computer program comprising:

at least one insertion point marker capable of maintaining the coordinates of an insertion point

on a display, the insertion point representing a location on the display where a user desires to provide input;

a speech recognition routine capable of decoding a speech signal; and

a meter generation routine capable of displaying a meter near the insertion point based on the insertion point marker, the meter being indicative of an amount of a speech signal that has been decoded by the speech recognition routine.

18. The computer program of claim 17 wherein the meter generation routine further comprises:

a microphone state variable having a value that is indicative of whether a microphone is active; and

an active microphone display routine, capable of displaying an indication that the microphone is active near the insertion point.

19. The computer program of claim 17 wherein the meter generation routine further comprises a transform routine capable of transforming a digital value into a set of coordinates for a shape on the display, the digital value being indicative of the magnitude of a portion of a speech signal.

20. The computer program of claim 17 wherein the speech recognition routine is capable of decoding a speech signal into a set of sub-words.

21. A method in a computer system for generating images on a display device, the method comprising:

receiving a speech input signal that is indicative of human speech;

displaying a volume meter that is indicative of the magnitude of at least a portion of the speech input signal; and

displaying a progress meter close to the volume meter on the display so that a user can perceive both the progress meter and the volume meter without substantially moving their eyes, the progress meter quantitatively indicating the amount of progress of a speech recognition system in decoding the speech input signal.

22. The method of claim 21 wherein displaying a volume meter comprises:

storing digital values representing the magnitudes of different respective portions of the speech signal;

accessing the stored digital values;

displaying a separate token for each separate digital value that is accessed.

23. The method of claim 22 wherein displaying a separate token comprises:

displaying a meter portion of the token, the meter portion's size being positively related to the magnitude of the speech signal such that higher magnitude portions of the speech signal have larger meter portions; and

displaying a background portion of the token, the background portion's size being negatively related to the magnitude of the speech signal such that

higher magnitude portions of the speech signal have smaller background portions.

24. The method of claim 23 wherein displaying the meter portion comprises:

transforming a digital value representing the magnitude of a portion of the speech signal to produce a transform value, the range between the smallest and largest transform value being less than the range between the smallest and largest digital value;

dividing the transform value by a maximum meter value to produce a meter ratio; and

determining the height of at least a portion of the meter portion using the meter ratio and a full meter height.

25. The method of claim 24 wherein determining the height of at least a portion of the meter portion comprises multiplying the meter ratio by the full meter height.

26. The method of claim 24 wherein determining the height of at least a portion of the meter portion comprises:

determining if the meter ratio is greater than a base ratio and if the meter ratio is greater than the base ratio performing steps comprising:

multiplying the base ratio by the full meter height to determine the height of a base block of the meter portion;

subtracting the base ratio from the meter ratio to produce an excess ratio;

using the excess ratio and the full meter height to determine a height of a second block of the meter portion.

27. The method of claim 26 wherein using the excess ratio and the full meter height to determine a height of a second block comprises:

determining if the excess ratio exceeds an intermediate ratio and if the excess ratio exceeds the intermediate ratio performing steps comprising:

multiplying the intermediate ratio by the full meter height to produce the height of the second block;

subtracting the intermediate ratio from the excess ratio to produce a remainder ratio; and

multiplying the remainder ratio by the full meter height to produce a height for a top block of the meter portion.

28. The method of claim 24 wherein the speech recognizer decodes the speech input signal by converting frames of the speech input signal into sub-words and wherein displaying a token comprises:

dividing the number of the last frame decoded by the speech recognizer by a total number of frames that form the speech input signal to produce a progress ratio;

multiplying the progress ratio by a full meter width to produce a progress width;

dividing the progress width by a token width to produce an affected number of tokens; and

for each of the affected number of tokens, setting the color of at least a portion of each token so that it is different from the color of other tokens.

29. A computer program designed to operate in a computer system having a display, the computer program comprising:

a volume meter portion capable of displaying a volume meter on the display that is indicative of the volume of a human speech signal;

a speech recognition portion that is capable of converting the human speech signal into a set of sub-words; and

a progress meter portion capable of displaying a progress meter on the display proximate the volume meter, the progress meter being quantitatively indicative of the amount of progress of the speech recognition portion in converting the human speech signal.

30. The computer program of claim 29 wherein the volume meter portion comprises:

meter size program code capable of determining a maximum dimension for the volume meter;

volume ratio program code capable of calculating a volume ratio that is defined as a magnitude value associated with the human speech signal over a maximum possible magnitude value; and

volume token program code capable of generating a volume token on the display that has a size that is determined from the volume ratio and the maximum dimension for the volume meter.

31. The computer program of claim 30 wherein the volume token program code comprises:

positive relation program code capable of generating a foreground portion of the volume token that is larger for higher magnitude values associated with the human speech signal; and

negative relation program code capable of generating a background portion of the volume token that is smaller for higher magnitude values associated with the human speech signal.

32. The computer program of claim 31 wherein the progress meter portion comprises:

progress ratio program code capable of dividing a frame number representing the last frame of the human speech signal converted by the speech recognition system by a total frame number representing the total number of frames found in the human speech signal to produce a progress ratio;

meter dimension program code capable of determining a maximum dimension for the progress meter; and

progress dimension program code capable of multiplying the progress ratio by the maximum dimension for the progress meter to produce a progress dimension.

33. The computer program of claim 32 wherein the volume token program code generates a volume token based in part on the progress dimension.

Appendix B

Cited References:

VanBuskirk et al. - U.S. Patent No. 6,075,534 - June 13, 2000

Tannenbaum - U.S. Patent No. 6,233,560 - May 15, 2001

Rozak et al. - U.S. Patent No. 5,864,815 - January 26, 1999

French-St. George et al. - U.S. Patent No. 6,018,711 - January 25, 2000

Appendix C

Table of Cases:

In re Lee, 277 F.3d 1361, 61 U.S.P.Q.2d 1430(Fed. Cir. 2002)

In re Rouffet, 149 F.3d 1350, 47 U.S.P.Q.2d 1453 (Fed. Cir. 1998)

APPENDIX D

(copies of references cited)

US005864815A

United States Patent [19]

Rozak et al.

[11] Patent Number: 5,864,815

[45] Date of Patent: Jan. 26, 1999

[54] METHOD AND SYSTEM FOR DISPLAYING SPEECH RECOGNITION STATUS INFORMATION IN A VISUAL NOTIFICATION AREA

[75] Inventors: **Michael J. Rozak**, Issaquah; **Juha P. Salin**, Bellevue; **James H. Spoltman**, Snohomish; **Ronald A. Belgau**, Port Angeles, all of Wash.

[73] Assignee: **Microsoft Corporation**, Redmond, Wash.

[21] Appl. No.: **508,757**

[22] Filed: **Jul. 31, 1995**

[51] **Int. Cl.**[6] **G10L 5/06**

[52] **U.S. Cl.** **704/275**; 704/270; 704/231; 395/326; 395/348

[58] **Field of Search** 395/2.84, 2.52, 395/2.43, 2.79, 345–348, 340; 364/513; 340/721

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,509,133	4/1985	Monbaron et al.	364/513.5
4,677,569	6/1987	Nakano et al.	364/513.5
4,704,696	11/1987	Reimer et al.	364/513.5
4,766,529	8/1988	Nakano et al.	364/513.5
4,776,016	10/1988	Hansen	381/42
4,783,803	11/1988	Baker et al.	381/42
4,827,520	5/1989	Zeinstra	381/43
4,864,623	9/1989	Van Nes et al.	381/43
4,866,778	9/1989	Baker	381/43
4,903,012	2/1990	Ohuchi	340/709
4,931,950	6/1990	Isle et al.	364/513
4,987,411	1/1991	Ishigami	340/709
5,027,406	6/1991	Roberts et al.	381/43
5,075,675	12/1991	Barker et al.	340/721
5,157,384	10/1992	Greanias et al.	340/706
5,231,670	7/1993	Goldhor et al.	381/43
5,231,691	7/1993	Yasuda	395/2
5,255,341	10/1993	Nakajima	395/2
5,265,202	11/1993	Krueger et al.	395/158
5,357,596	10/1994	Takebayashi et al.	395/2.84
5,384,910	1/1995	Torres	395/156
5,386,494	1/1995	White	395/2.84
5,404,458	4/1995	Zetts	395/275
5,408,582	4/1995	Colier	395/2.52
5,425,141	6/1995	Gedye	395/157
5,444,617	8/1995	Merialdo	364/419.1
5,448,695	9/1995	Douglas et al.	395/155
5,465,378	11/1995	Duensing et al.	395/800
5,519,809	5/1996	Husseiny et al.	395/2.84
5,524,051	6/1996	Ryan	380/9
5,566,248	10/1996	Ulrich	382/187
5,590,264	12/1996	Keane et al.	395/340
5,602,963	2/1997	Bissonnette et al.	395/2.84
5,602,981	2/1997	Hargrove	395/352
5,617,526	4/1997	Oran et al.	395/326

OTHER PUBLICATIONS

Microsoft Corporation, "User's Guide—Microsoft® Word—The World's Most Popular Word Processor," Version 6.0, pp. 72–73, 1993–1994.

Kurzweil, "Kurzweil Brings Voice Dictation to Windows," *Byte*, 19(8):48, Aug., 1994.

Meisel, William S., "Talk to Your Computer," *Byte*, 18(11): 113, Oct., 1993.

Schmandt et al., "Augmenting a Window System with Speech Input," *Computer*, 23(8): 50–56, Aug., 1990.

Borgen et al., "Voice Navigator—Owner's Guide," *Articulate Systems, Inc.*, pp. 4, 11–12, 25–27, and 43–44, 1990.

"Integrated Audio–Graphics User Interface," *IBM Technical Disclosure Bulletin*, 33(11): 368–371, Apr. 1991.

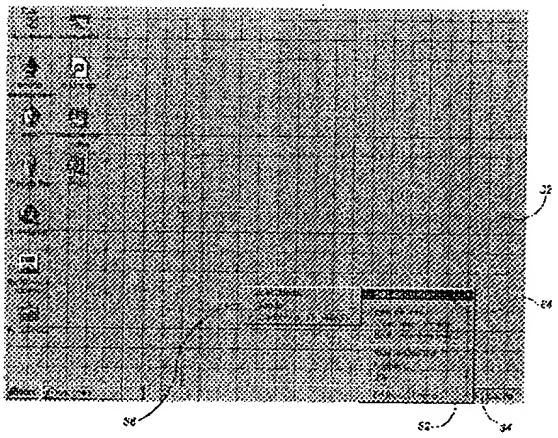
"Software User's Guide," *Microsoft Windows Sound System Version 2.0*, pp. 28–70, 1993.

Primary Examiner—Richemond Dorvil
Assistant Examiner—Vijay B. Chawan
Attorney, Agent, or Firm—Westman, Champlin & Kelly, P.A.

[57] **ABSTRACT**

A speech recognition system provides a user with graphical and textual feedback. The textual feedback is displayed in windows but occupies little of the available display space and are displayed only for a short period of time. The graphical feedback is displayed in a designated notification area and does not obscure any other displayed items. The feedback provided by the speech recognition system may indicate a current mode of operation of the speech recognition system as well as a state of processing of audio input by the speech recognition system.

38 Claims, 14 Drawing Sheets



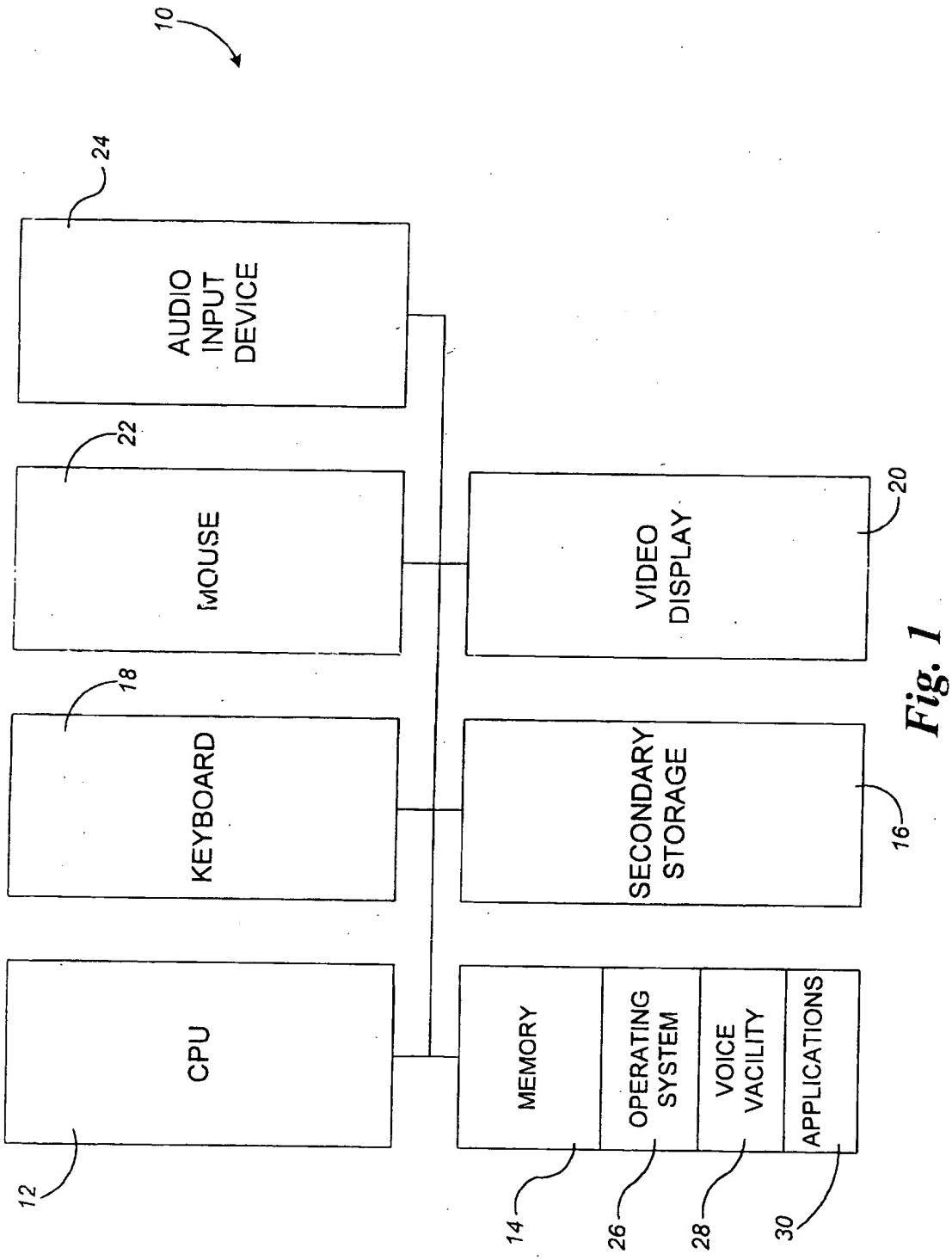

10
24
AUDIO INPUT DEVICE
22
MOUSE
18
KEYBOARD
CPU
12
VIDEO DISPLAY
20
SECONDARY STORAGE
16
MEMORY
OPERATING SYSTEM
VOICE VACILITY
APPLICATIONS
14
26
28
30

Fig. 1

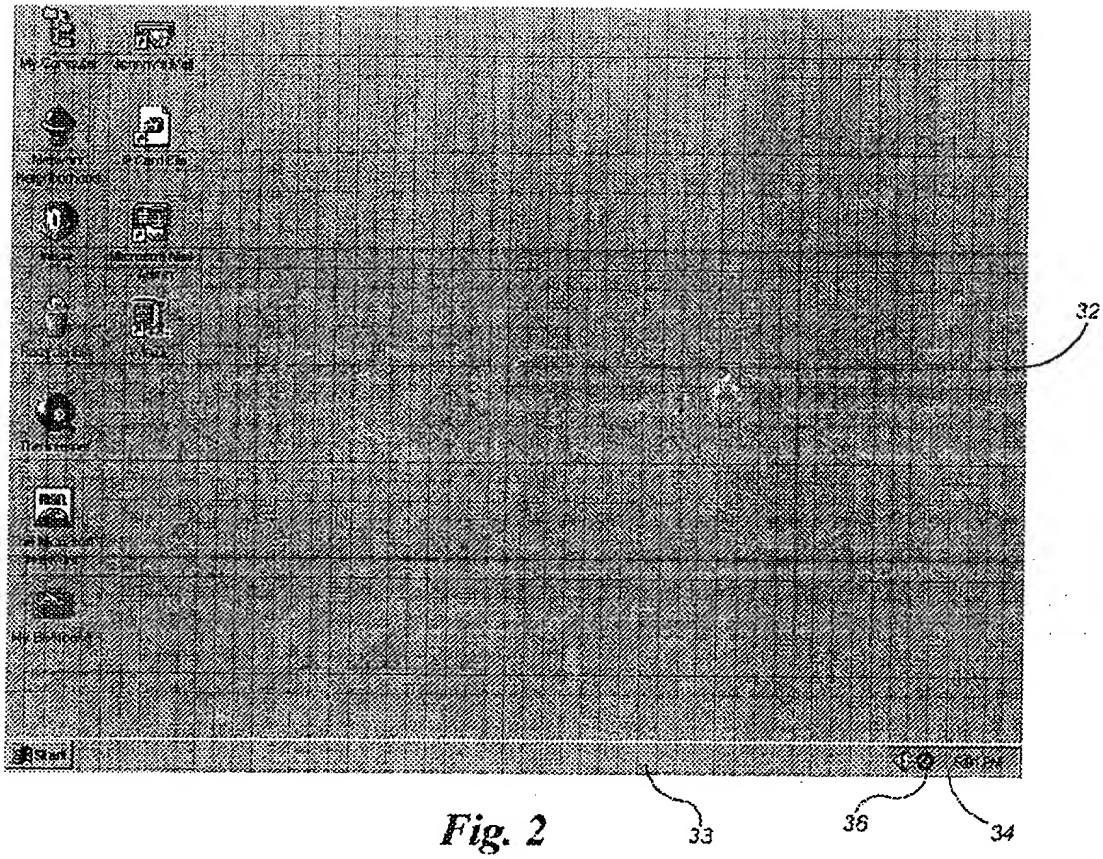


Fig. 2

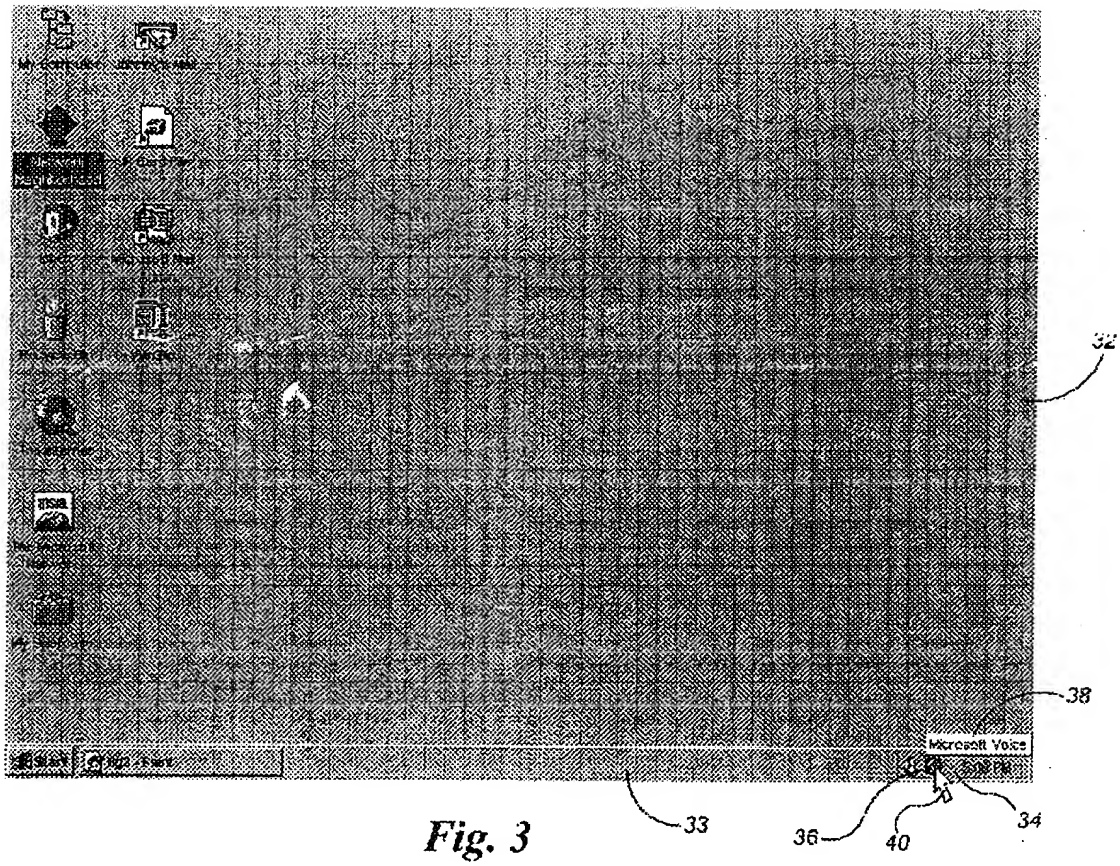


Fig. 3

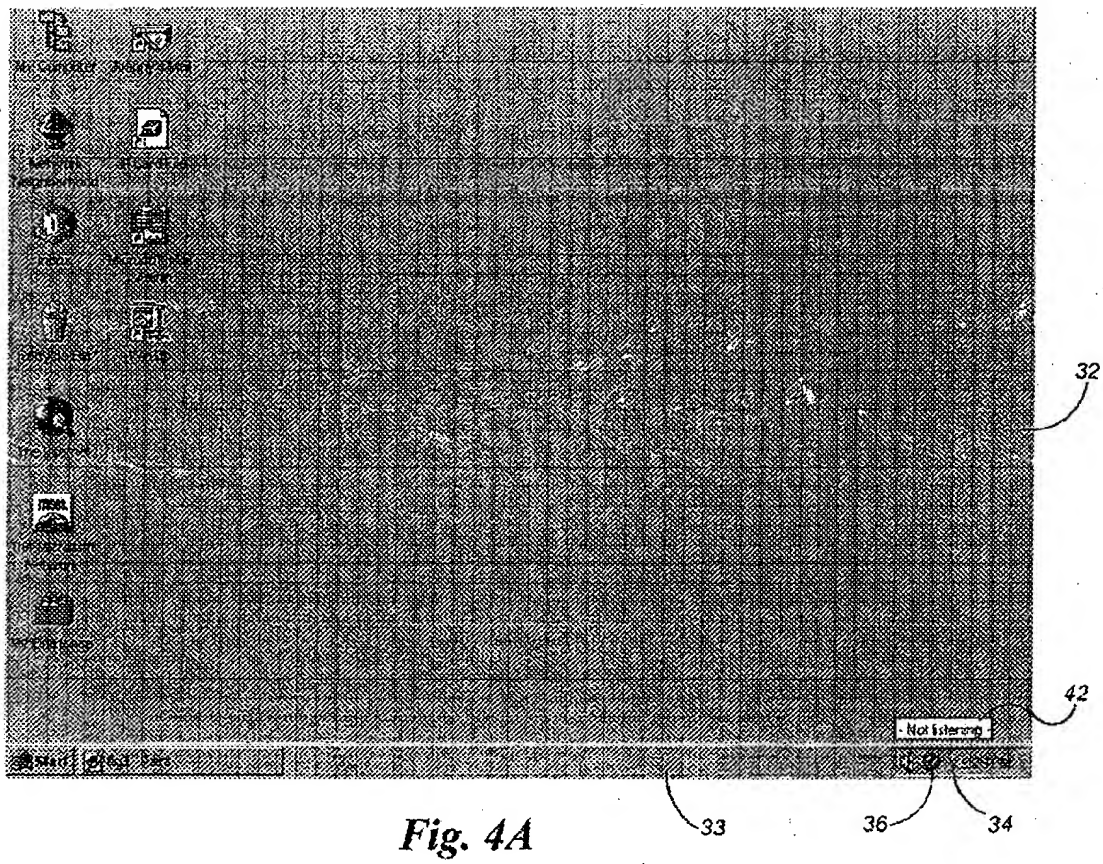


Fig. 4A

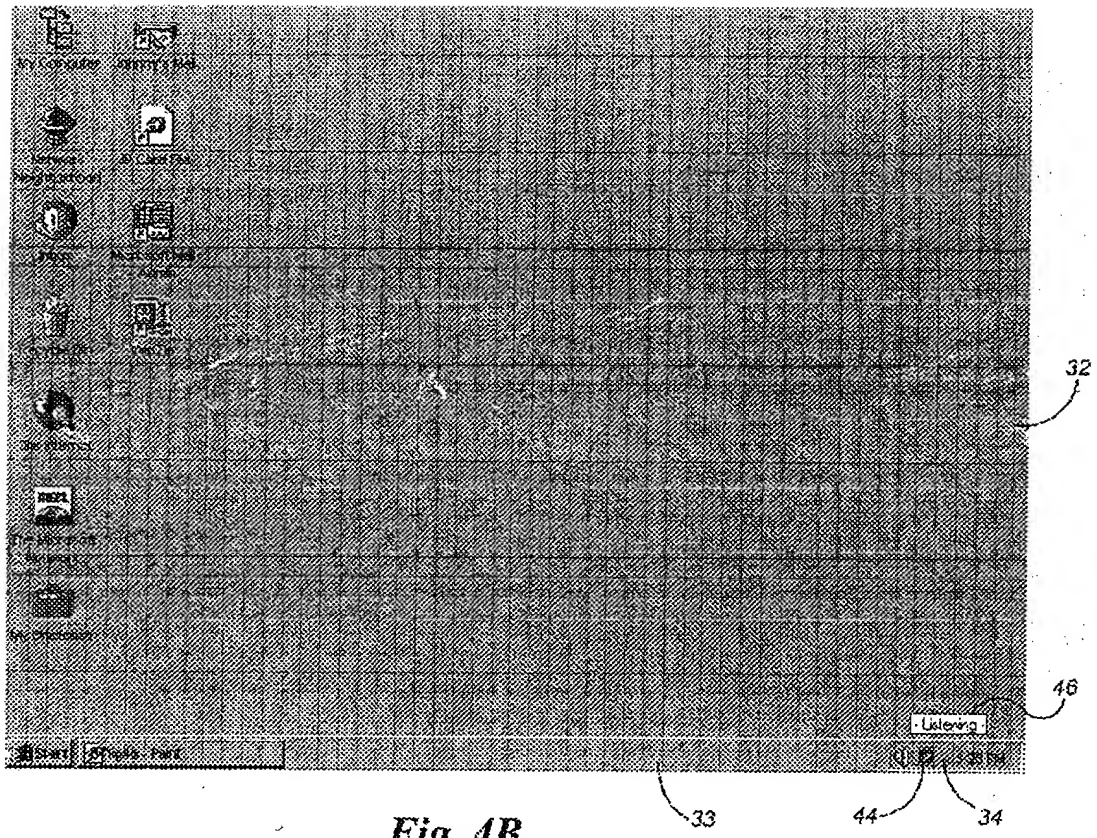


Fig. 4B

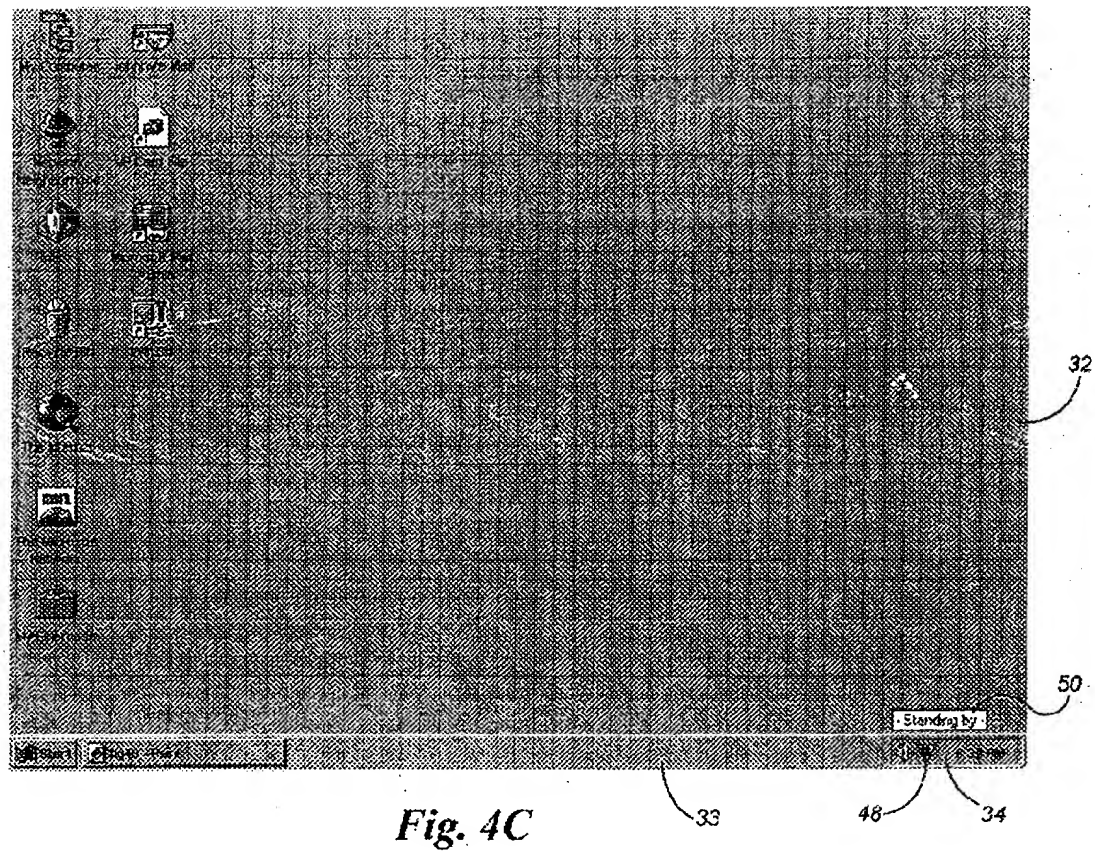


Fig. 4C

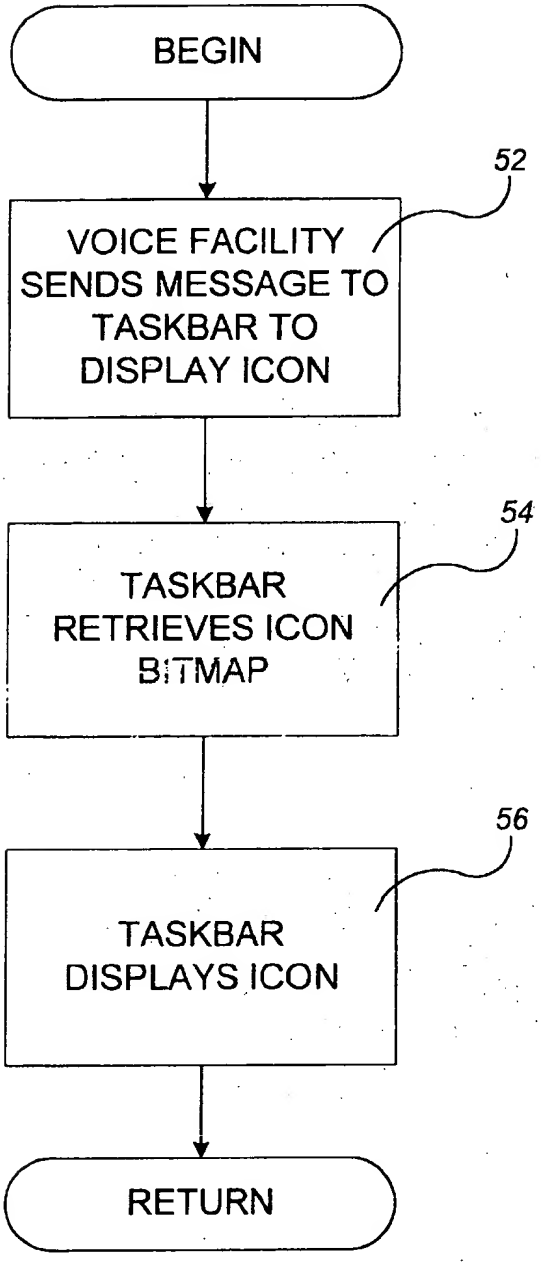

BEGIN
52
VOICE FACILITY SENDS MESSAGE TO TASKBAR TO DISPLAY ICON
54
TASKBAR RETRIEVES ICON BITMAP
56
TASKBAR DISPLAYS ICON
RETURN

Fig. 5

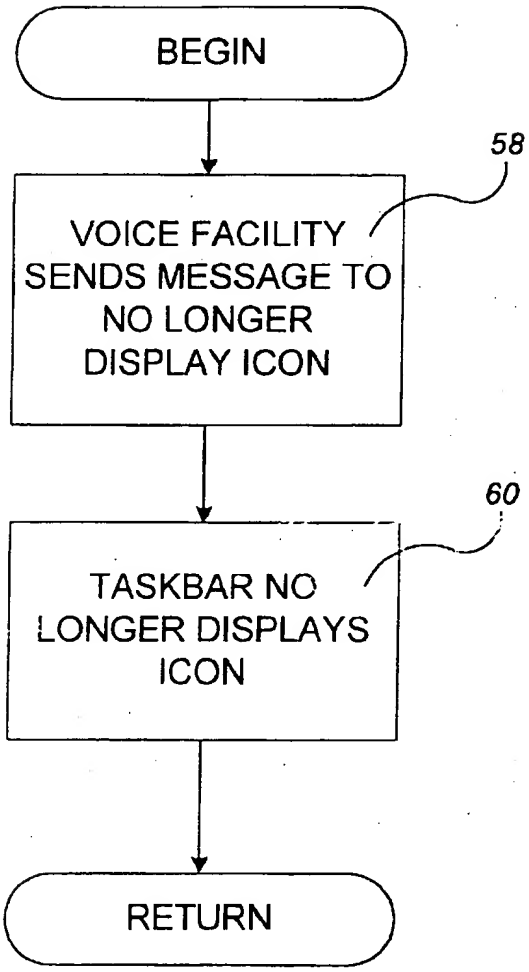

BEGIN
58
VOICE FACILITY SENDS MESSAGE TO NO LONGER DISPLAY ICON
60
TASKBAR NO LONGER DISPLAYS ICON
RETURN

Fig. 6

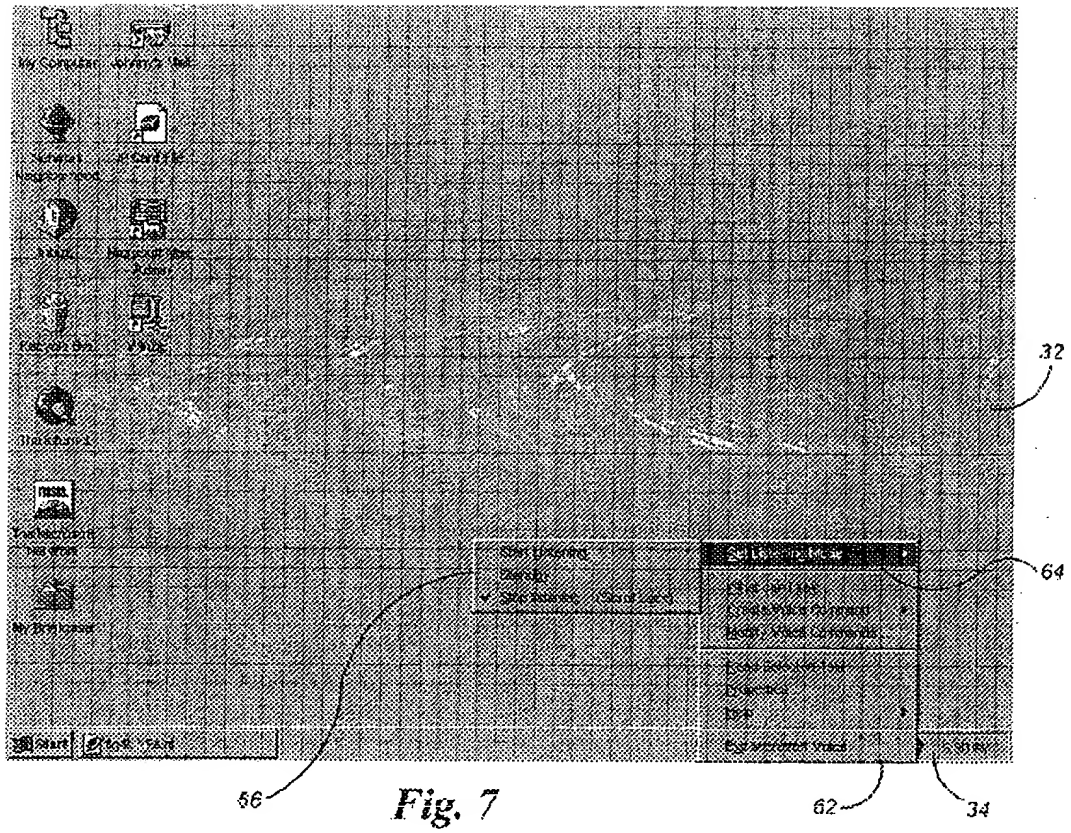


Fig. 7

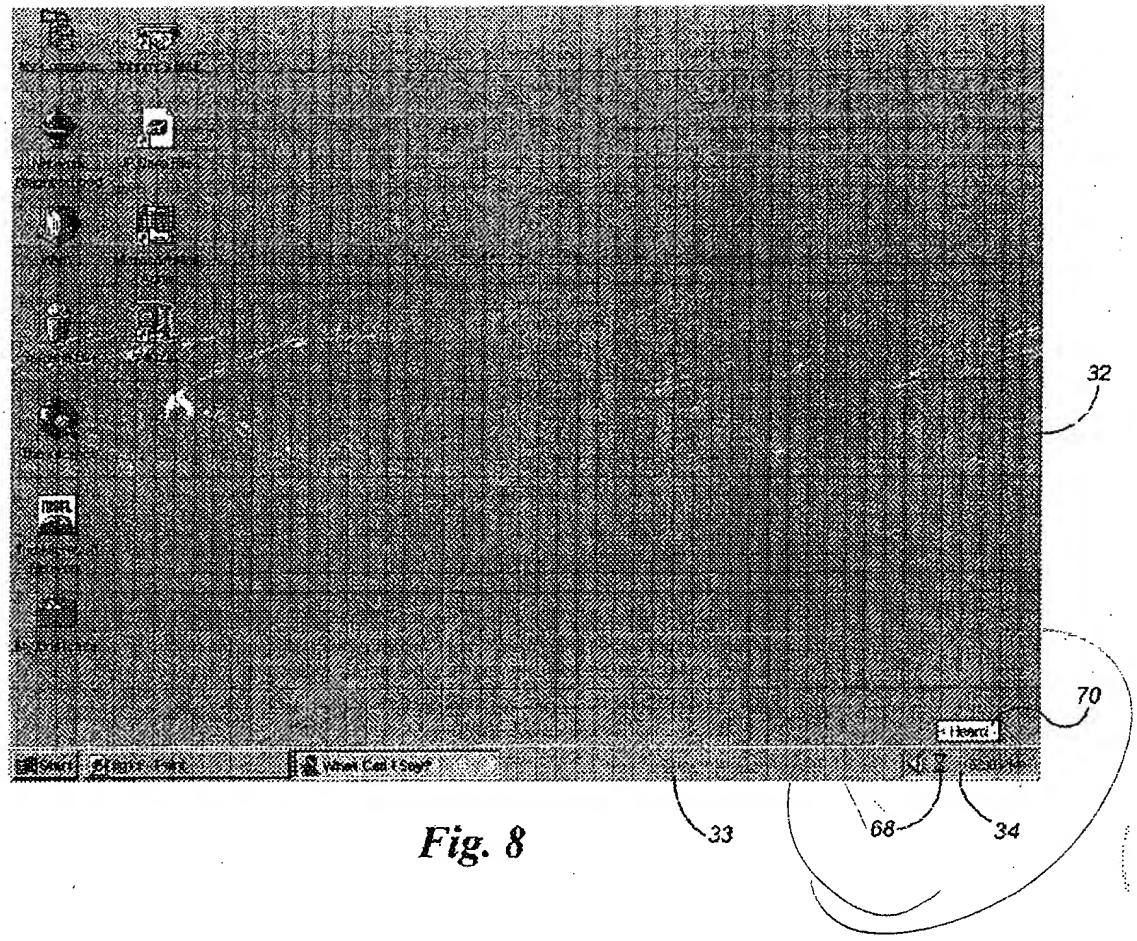


Fig. 8

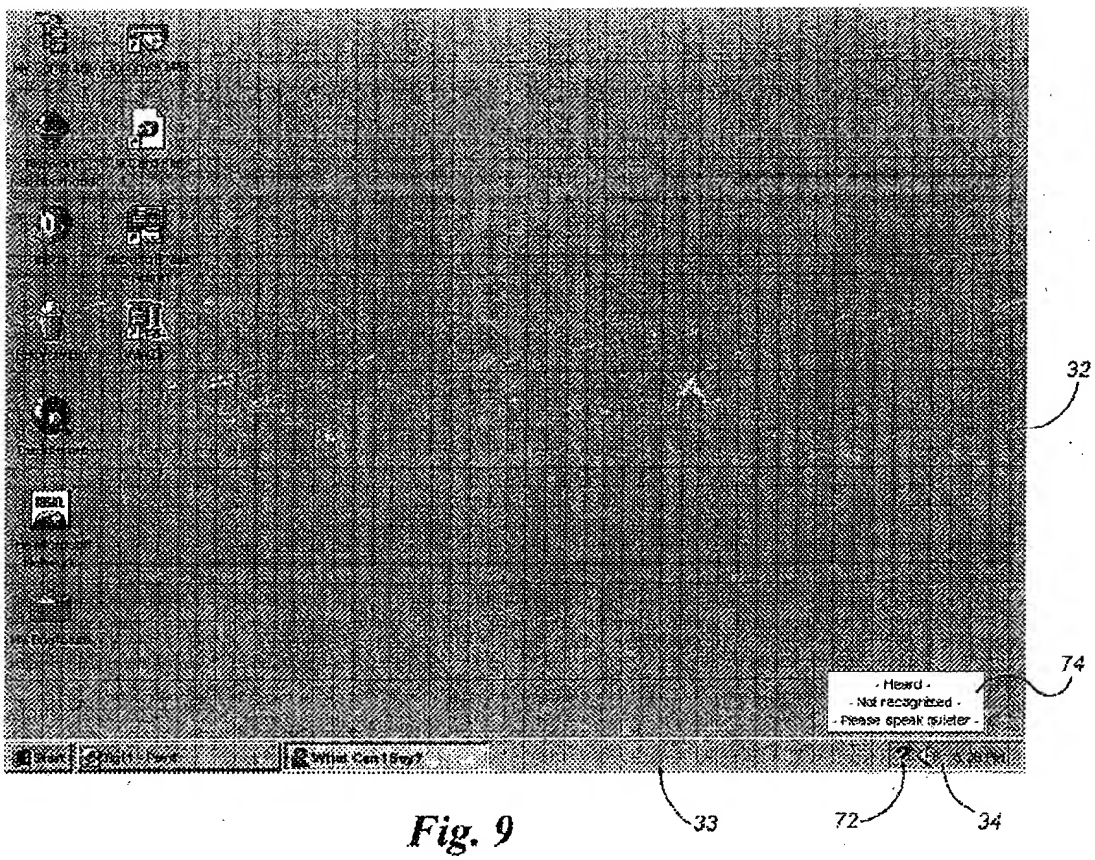


Fig. 9

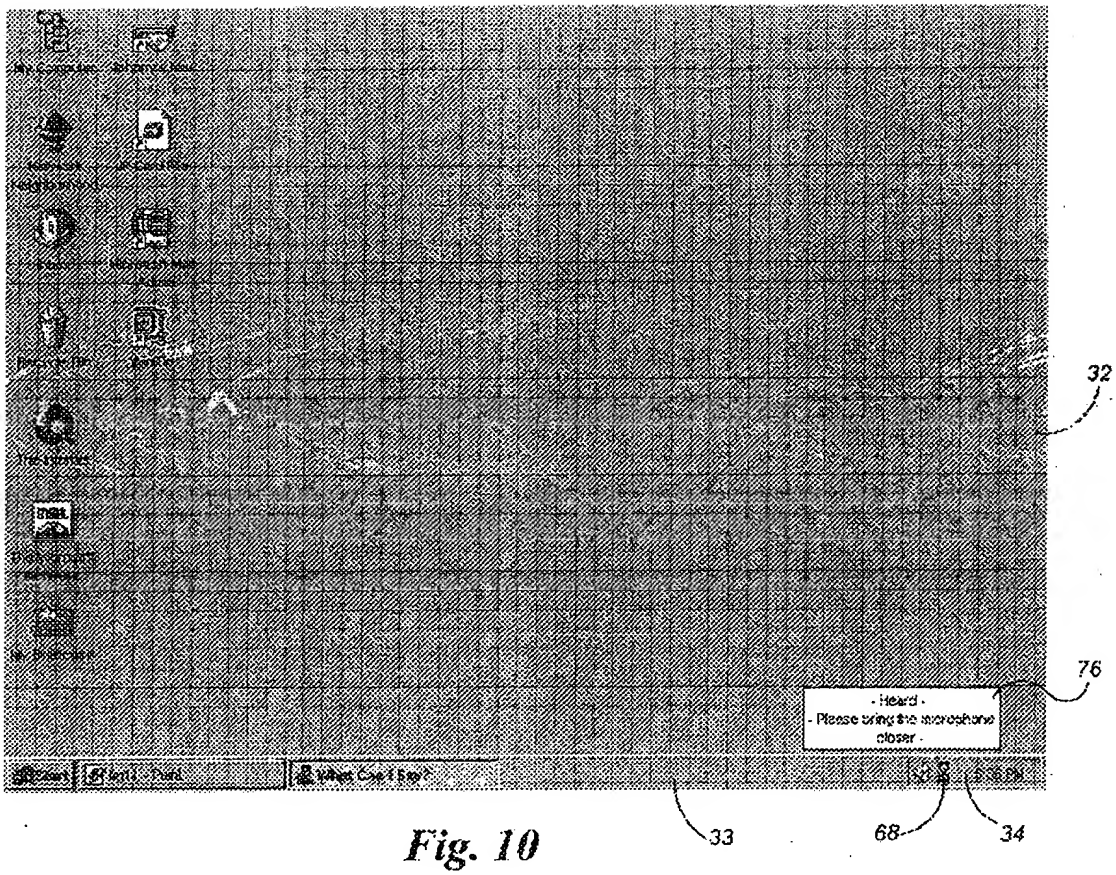


Fig. 10

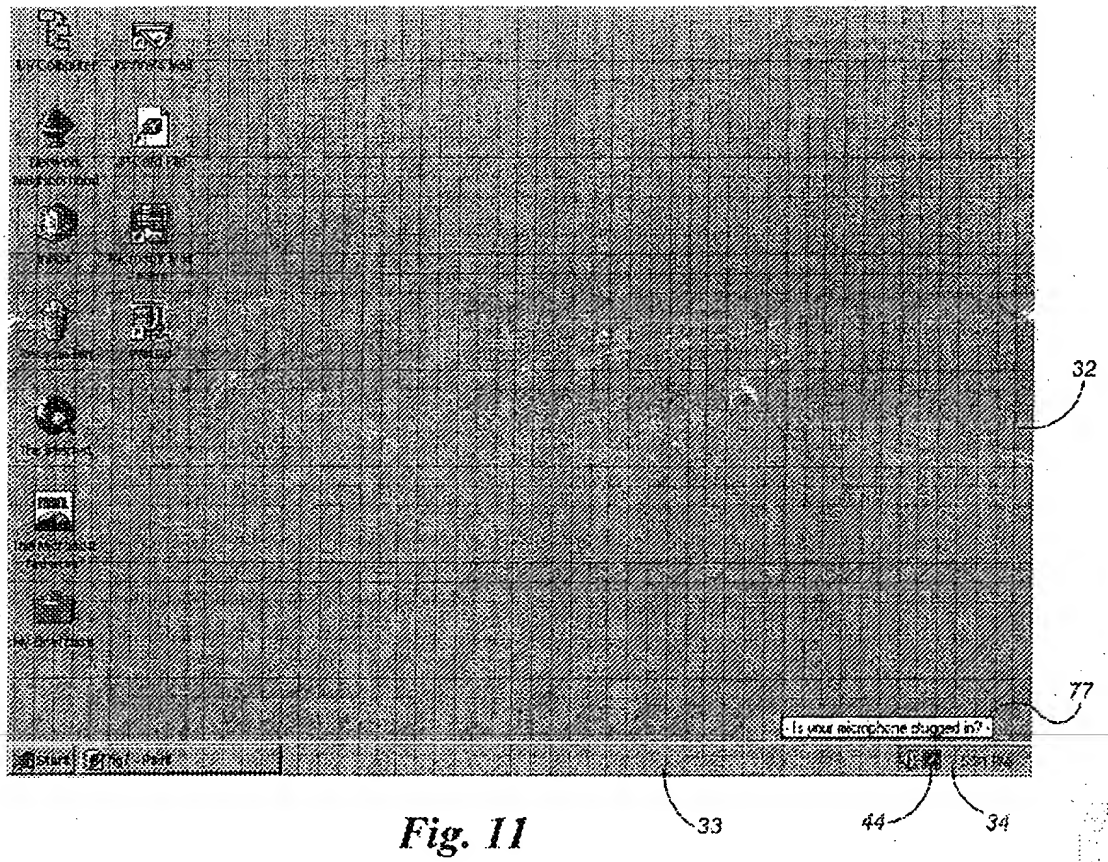


Fig. 11

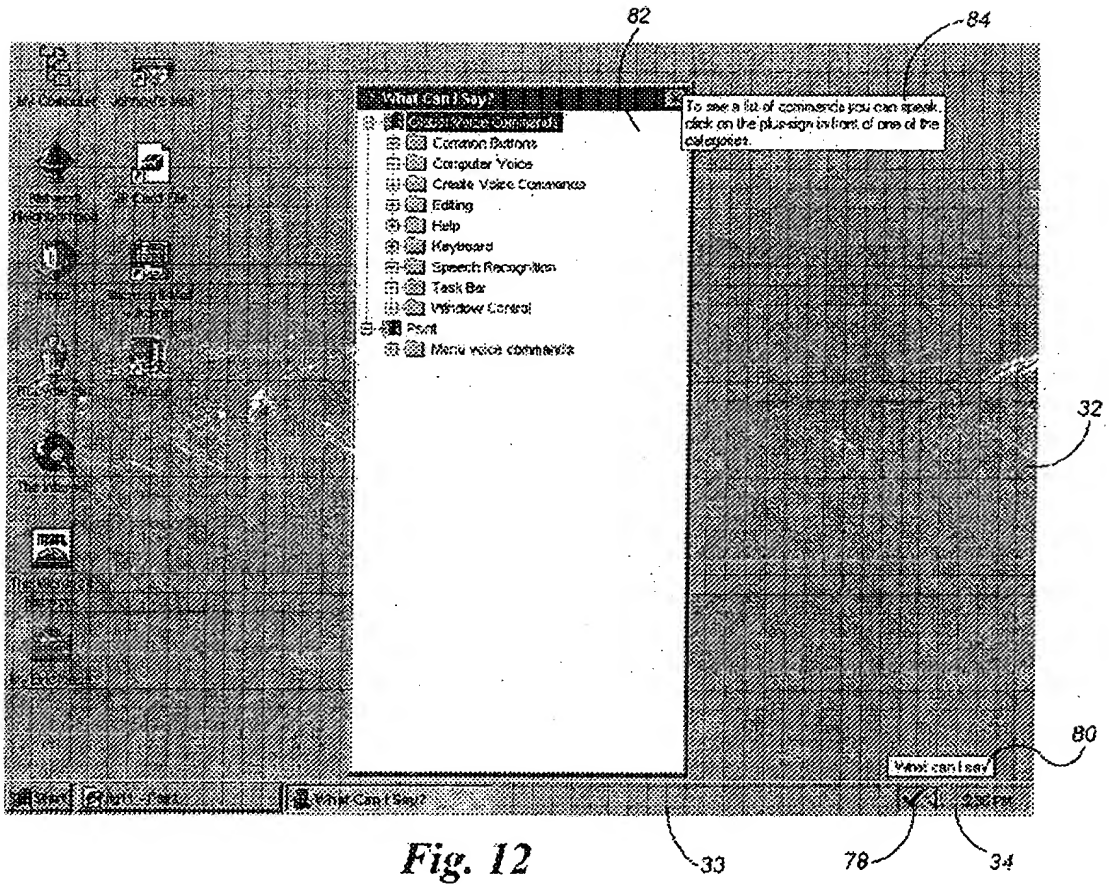

82
84
What Can I Say?
To see a list of commands you can speak, click on the plus sign in front of one of the categories.
Common Buttons
Computer Voice
Create Voice Commands
Editing
Help
Keyboard
Speech Recognition
Task Bar
Window Control
Print
Menu voice commands
32
80
What can I say
33
78
34

Fig. 12

METHOD AND SYSTEM FOR DISPLAYING SPEECH RECOGNITION STATUS INFORMATION IN A VISUAL NOTIFICATION AREA

TECHNICAL FIELD

The present invention relates generally to computer systems and more particularly to speech recognition in computer systems.

BACKGROUND OF THE INVENTION

Speech recognition systems provide facilities for recognizing components of speech in audio input. When a speaker speaks into an audio input device, the speech recognition systems process the audio input to recognize speech components. One difficulty encountered with such speech recognition systems is the difficulty in informing the user of how the audio input was interpreted and whether the audio input was properly received and processed. Certain conventional speech recognition systems have attempted to address this problem by providing a user interface, such as a speech recognition window that provides feedback to the user about the progress and results of speech recognition. Unfortunately, such windows often obscure substantial portions of the available display space on the video display. In addition, such windows are typically displayed as long as the speech recognition system is running so that the windows continue to obscure portions of the video display while the speech recognition system is running.

SUMMARY OF THE INVENTION

The present invention overcomes the limitations of the prior art by displaying feedback to the user in a fashion that minimizes the extent and duration to which other displayed objects are obscured by the displayed feedback. In accordance with a first aspect of the present invention a method is practiced in a computer system that has a video display and an audio input device for receiving audio input. The computer system also includes a speech recognizer for recognizing components of speech in the audio input. The speech recognizer operates in one of a number of different modes of operation. For this method, a graphical user interface that includes a designated notification area for displaying graphical notifications is displayed on the video display. The current mode of operation of the speech recognizer is determined, and a graphical notification of the current mode of operation of the speech recognizer is displayed in the designated notification area.

In accordance with another aspect of the present invention, audio input is received from the user through an audio input device. The received audio input is processed by a speech recognizer to attempt to recognize components of speech. Textual feedback regarding the processing of the received audio input is displayed in a window that is displayed for only a predetermined period of time, such as a few seconds.

In accordance with a further embodiment of the present invention, the computer system includes a video display for displaying video images and an audio input device for receiving audio input from a user. The computer system additionally includes a speech recognizer for recognizing components of speech in audio input that is received by the audio input device. The computer system further includes a component for providing a graphical notification area on the video display for displaying graphical notifications to the user. The computer system also includes a facility for displaying a graphical notification in the graphical notification area that identifies a current mode of operation of the speech recognizer.

In accordance with yet another aspect of the present invention a computer system includes a video display, an audio input device and a speech recognizer. The computer system, likewise, includes a facility for displaying textual feedback to the user regarding processing of the received audio input by the speech recognizer in a window that is only displayed for a predetermined period of time.

In accordance with yet another aspect of the present invention, a computer-readable storage medium is designed for use in a computer system that includes a video display and an audio input device for receiving audio input from a user. The computer-readable storage medium stores a speech recognizer for recognizing components of speech in the audio input that is received by the audio input device. The computer-readable storage media also stores a first facility for providing a graphical notification area on the video display in which to display graphical notifications. The computer-readable storage medium further stores a second facility for displaying graphical notifications in the graphical notification area regarding the speech recognizer.

BRIEF DESCRIPTION OF THE DRAWINGS

A preferred embodiment of the present invention will be described below with reference to the following drawings.

FIG. **1** is a block diagram of a computer system that is suitable for practicing the preferred embodiment of the present invention.

FIG. **2** is an example illustrating the display of a state icon for the voice facility in the visual notification area in accordance with the preferred embodiment of the present invention.

FIG. **3** illustrates the display of a tool tip when a mouse cursor points at a state icon for the voice facility in accordance with the preferred embodiment of the present invention.

FIG. **4A** illustrates the display of graphical and textual information in response to the changing the state of the voice facility to the not listening state.

FIG. **4B** illustrates the display of graphical and textual information in response to the changing of the state of the voice facility to the listening state.

FIG. **4C** illustrates the display of graphical and textual information in response to the changing of the state of the voice facility to the standing-by state.

FIG. **5** is a flow chart illustrating the steps that are performed to display an icon by the voice facility in the graphical notification area.

FIG. **6** is a flow chart illustrating the steps that are performed to remove an icon from the graphical notification area.

FIG. **7** illustrates a menu sequence that is displayed to enable a user to change current state of the voice facility per the preferred embodiment of the present invention.

FIG. **8** illustrates the display of a tool tip that indicates that audio input was heard by the voice facility and the display of an icon that indicates that the voice facility is processing the audio input.

FIG. **9** illustrates a tool tip that indicates that audio input was heard, not recognized and too loud and also illustrates an icon that indicates the audio input was not recognized by the voice facility.

FIG. **10** illustrates the display of a tool tip that is generated when the signal to noise ratio of received audio input is below an acceptable level.

FIG. **11** illustrates the display of a tool tip that is displayed when the voice facility receives no audio input signal when in the listening state.

FIG. **12** illustrates an example of the appearance of the video display when a voice command is recognized by the voice facility.

DETAILED DESCRIPTION OF THE INVENTION

The preferred embodiment of the present invention provides a method and system for providing speech recognition status information to a user in a visual notification area. The information is displayed in a fashion that occupies very little of the available display space on the video display and much of the information is displayed in a transitory or temporary fashion so that any obstruction that is caused by the displaying of the status information is only temporary.

FIG. **1** shows a block diagram of a computer system **10** that is suitable for practicing the preferred embodiment of the present invention. Although the computer system **10** of FIG. 1 includes only a single processor, those skilled in the art will appreciate that the present invention may also be practiced in systems that employ multiple processors, including distributed systems. The computer system **10** of FIG. 1 includes a central processing unit (CPU) **10** that is accessed to a primary memory **14** and secondary storage **16**. The secondary storage **16** may take many forms including a hard disk drive. The computer system **10** also includes a keyboard **18**, a video display **20**, a mouse **22** and an audio input device **24**. The audio input device **24** may take the form of a microphone that is coupled to a sound board. The computer system **10** need not include all of these peripheral devices; rather, systems lacking one or more of these peripheral devices may be used to practice the present invention.

The memory **14** holds a copy of any operating system **26**, voice facility **28** and application programs **30**. For illustrative purposes in the discussion below, the operating system **26** is assumed to be the "MICROSOFT" "WINDOWS" '95 operating system sold by Microsoft Corporation of Redmond, Wash. The voice facility **28** includes a speech recognition engine for recognizing speech components in audio input and a voice command facility for responding to voice commands that are recognized by the speech recognizer. The application programs **30** may call upon facilities of the operating system **26** and the voice facility **28**.

The preferred embodiment of the present invention relays status information regarding the current state of the voice facility **28**. One way in which status information is relayed to user is by displaying icons within a visual notification area. As is shown in FIG. **2**, the operating system **26** has a taskbar **33** in which buttons associated with active tasks are displayed. A predesignated portion of the taskbar **33** is designated as a taskbar visual notification area **34** in which application programs **30**, the voice facility **28** and the operating system **26** may display icons to convey status information to a user. The taskbar visual notification area is described in more detail in co-pending application entitled "System Provided Visual Notification Area", Ser. No. 08/355,398, U.S. Pat. No. 5,617,526 which was filed on Dec. 13, 1994 and is assigned to a common assignee with the present application. The co-pending application is incorporated by reference herein.

In order to understand status information that is displayed by the voice facility **28**, it is helpful to examine the different states which the voice facility may assume. The voice facility **28** may assume a listening state wherein the voice facility listens for audio input and processes all audio input. The voice facility **28** may also assume a not listening state wherein no audio input received by audio input device **24** is processed. Lastly, the voice facility **28** may assume a standing-by state wherein the voice facility analyses audio input that is received by the audio input device **24** but only is responsive to selected voice commands, including voice commands that are prefaced by a trigger phrase, such as "computer," and voice commands, such as "start listening," that change the state of the voice facility.

The current state of the voice facility **28** is reflected in an icon that is displayed within the taskbar visual notification area **34** (FIG. **2**). FIG. **2** shows an example of the not listening state icon **36**. The presence of the not listening icon **36** in the taskbar visual notification area **34** indicates that the voice facility is not currently listening to audio input from the audio input device **24**. One of the strengths of displaying state icons in the visual notification area **34** is that it provides a known and easily recognizable location for displaying such information. Moreover, the icons displayed within the visual notification area **34** do not occupy any of the virtual desktop surface **32** and, hence, do not obscure other displayed items.

If a user uses the mouse **22** (FIG. 1) or another input device to position a mouse cursor **40** (FIG. **3**) over a state icon, a tool tip will be displayed. FIG. **3** shows an example wherein a mouse cursor **40** is positioned to point at the not listening state icon **36** which is displayed in the taskbar visual notification area **34**. A tool tip **38** is displayed to indicate that the icon is associated with "MICROSOFT VOICE." The tool tip **38** is a small window that is displayed in proximity to the icon and that contains text. The tool tip is implemented as a clipped rectangle that displays text provided by the application associated with the icon. Thus, the voice facility **28** provides the text that is to be displayed within the clipped rectangle for the tool tip **38**. The tool tip **38** is only displayed for a predetermined period of time, such as 2 or 3 seconds. After the expiration of the predetermined time, the tool tip **38** is no longer displayed and, thus, no longer occupies real estate on the virtual desktop **32**.

A separate icon is provided for each of the three states. FIG. **4B** shows the listening state icon **44**, and FIG. 4C shows the standing-by state icon **48**. FIGS. **3** and **4A** show the non-listening state icon. When the voice facility state is changed, the state icon for the new state is displayed within the taskbar notification area **34**. In addition, a tool tip that identifies the new state is displayed. FIG. **4A** shows the tool tip **42** that is displayed when the state changes to the not listening state. FIG. **4B** shows a tool tip **46** that is displayed when the state is changed to the listening state, and FIG. 4C shows the tool tip **50** that is displayed when the state is changed to the standing-by state. The voice facility **28** is responsible for providing the text for such tool tips and initiating display of the tool tips.

FIG. **5** is a flow chart showing the steps that are performed in order for voice facility state icons to be displayed in the taskbar visual notification area. Initially, the voice facility **28** sends a message to the taskbar to display the appropriate state icon (step **52**). This message includes identification information that enables the taskbar to identify where the bitmap for the icon is stored. The taskbar then retrieves the icon (step **54**). Once the taskbar has retrieved the icon bitmap, the taskbar displays the icon using the bitmap (step **56**).

The removal of a state icon so that it no longer is displayed in the taskbar visual notification area **34** is also

message-driven. First, the voice facility **28** sends a message to the taskbar to no longer display the state icon (step **58** in FIG. **6**). The taskbar receives this message and, in response, removes the icon from the taskbar visual notification area **34** (step **60**). If the voice facility is still active, the taskbar may then proceed to display a new state icon in the taskbar visual notification area.

The taskbar visual notification area **34** is a window. The taskbar knows where each icon is displayed within the taskbar visual notification area **34**. As such, when the mouse cursor **50** is positioned over one of the state icons, the taskbar knows that the mouse cursor is pointing at one of the icons and asks the voice facility **28** what to do in response to the mouse cursor pointing to the state icon. The voice facility **28** causes a tool tip to be displayed after the expiration of a predetermined time period if the mouse cursor remains pointing at the state icon the tool tip identified that the state icon is associated with the "MICROSOFT VOICE" application.

As mentioned above, a user has the ability to select a current state for the voice facility **28**. One manner in which the user may select the state of the voice facility **28** is through a state icon. If the user positions the mouse cursor **50** so that it points at the state icon that is displayed within the taskbar visual notification area **34** and depresses one of the mouse buttons a menu **62** (FIG. **7**) is displayed. One of the choices on the menu **62** is a "Set Listening Mode" menu option **64** that present a slide-out menu **66**. The slide-out menu **66** displays each of the available states and allows a user to select a state by positioning the mouse cursor on one of the states and releasing the mouse button. The current state is indicated by a check mark in the slide-out menu **66**. For the example shown in FIG. **7**, the voice facility **28** is currently in the not listening state. Alternatively, a dialog may be displayed rather than the slideout menu.

The voice facility **28** also provides the user with another type of status information. Specifically, the voice facility **28** provides the user with feedback regarding the audio input that has been entered by the user through the audio input device **24**. In response to audio input, tool tips are displayed, and the state icon may be changed to icons that indicate the state of processing by the voice facility **28**.

FIG. **8** shows an example of what is displayed to a user immediately after the user enters audio input that was sufficiently audible for the voice facility **28** to hear it. In particular, a tool tip **70** is displayed that indicates that the audio input was "Heard." In addition, the state icon is replaced with an icon **68** that indicates that the voice facility **28** is currently processing the audio input. If the voice facility does not recognize the voice commands that have been input as audio input by the user, the processing icon **68** is replaced with a question mark icon **72** as shown in FIG. **9**. The question mark icon **72** is an intuitive indication that the voice facility **28** does not recognize the audio input. In addition, the phrase "Not recognized" will be displayed within a tool tip **74**.

One of the reasons that the audio input may not be recognized is if the user is speaking too loudly into the audio input device **24**. The voice facility **28** determines that the audio input is too loud when the audio input exceeds a given threshold of loudness. In such a case, the tool tip **74** also includes the message "Please speak quieter." The audio input may also be too quiet. The voice facility **28** conveys this feedback by displaying the message "Please speak louder" in a tool tip. Similarly, when the signal to noise ratio of the received audio input is too low, the message "Please bring the microphone closer" is displayed in a tool tip **76**, such as shown in FIG. **10**. If no audio input signal at all is received, a tool tip **77** (FIG. **11**) that displays the message "Is your microphone plugged in?" is displayed.

The preferred embodiment of the present invention also provides feedback when the voice command input by a user is properly recognized. FIG. **12** shows an example wherein the voice command "What can I say?" is properly recognized. The recognition icon **78**, which resembles a check mark, is displayed within the taskbar visual notification area **34**. In addition, a tool tip **80** is displayed with the phrase "What can I say?" The voice facility **28** invokes the script that is associated with the command. In the example shown in FIG. **12** a window **82** is opened that provides a list of global voice commands. Furthermore, the voice facility **28** displays an additional tool tip **84** that explains operation of the hierarchical list that is displayed in the window **82**.

While the present invention has been described with reference to a preferred embodiment thereof, those skilled in the art will appreciate that various changes in form and detail may be made without departing from the intended scope of the invention as defined in the appended claims. For example, tool tips that differ from those displayed in the figures may be used to practice the present invention. Similarly, different state icons may be used to practice the present invention, and different voice facility states (other than those described) may be used to practice the present invention.

We claim:

1. In a computer system having a video display, an audio input device for receiving audio input from a user and a speech recognizer for recognizing components of speech in audio input, said speech recognizer operating in a number of different modes of operation, a method comprising the computer-implemented steps of:

providing a graphical user interface on the video display, said graphical user interface including a designated notification area for displaying graphical notifications;

determining a first current mode of operation of the speech recognizer;

displaying a first graphical notification of the first current mode of operation of the speech recognizer in the designated notification area;

determining a second current mode of operation of the speech recognizer; and

displaying a second graphical notification of the second current mode of the speech recognizer in the designated notification area, wherein the first notification differs from the second graphical notification.

2. The method of claim **1** wherein the first current mode of operation of the speech recognizer is determined to be a listening mode in which the speech recognizer is listening for and processing audio input received at the audio input device.

3. The method of claim **1** wherein the first current mode of operation of the speech recognizer is determined to be a non-listening mode in which the speech recognizer is not listening for or processing the audio input received at the audio input device.

4. The method of claim **1** wherein the first current mode of operation of the speech recognizer is determined to be a standing-by mode in which the speech recognizer is listening for and processing only selected audio input received at the audio input device.

5. The method of claim **1** wherein the step of displaying the graphical notification of the first current mode of opera-

tion of the speech recognizer in the designated notification area comprises displaying an icon that identifies the first current mode of operation of the speech recognizer in the designated notification area.

6. The method of claim 1 wherein the computer system runs an operating system and the operating system provides the graphical user interface.

7. The method of claim 1 further comprising the step of displaying text in a separate window that identifies the second current mode of operation.

8. The method of claim 7 further comprising the step of terminating the display of the text and the separate window after expiration of a predetermined time period.

9. The method of claim 1 wherein the computer system further comprises a mouse with at least one button, said mouse for manipulating a mouse cursor that points to locations on the video display and wherein the method further comprises the step of displaying a user interface element that enables the user to change the first current mode of operation of the speech recognizer in response to the mouse cursor pointing at the graphical notification and the user clicking the button on the mouse.

10. The method of claim 1 wherein the designated notification area does not obscure any other video output on the video display.

11. In a computer system having a video display, an audio input device for receiving audio input from a user and a speech recognizer for recognizing components of speech in audio input, a method comprising the computer-implemented steps of:

receiving audio input from the user through the audio input device;

processing the received audio input with the speech recognizer to attempt to recognize components of speech in the received audio input; and

displaying textual feedback regarding the processing of the received audio input by the speech recognizer in a tool tip window that is only displayed for a predetermined period of time.

12. The method of claim 11 wherein the displayed textual feedback indicates that the received audio input was recognized by the speech recognizer.

13. The method of claim 12 wherein the displayed textual feedback indicates what the received audio input was recognized as.

14. The method of claim 11 wherein the displayed textual feedback indicates that the received audio input was heard by the speech recognizer.

15. The method of claim 11 wherein the displayed textual feedback indicates that the received audio input was not recognized by the speech recognizer.

16. The method of claim 11 wherein the displayed textual feedback indicates that there are problems with the received audio input that make it difficult for the speech recognizer to process the received audio input.

17. The method of claim 11 further comprising the steps of:

providing a designated notification area on the video display for displaying graphical notifications by application programs;

displaying a graphical notification in the designated notification area to provide feedback to the user about processing of the received audio input by the speech recognizer.

18. The method of claim 17 wherein the displayed graphic notification indicates that the received audio input was recognized by the speech recognizer.

19. The method of claim 17 wherein the displayed graphical notification indicates that the received audio input was not recognized by the speech recognizer.

20. A computer system comprising:

a video display for displaying video images;

an audio input device for receiving audio input from a user;

a speech recognizer for recognizing components of speech in audio input received by the audio input device;

a component for providing a graphical notification area on the video display for displaying different types of graphical notifications to the user; and

a facility for displaying graphical notifications in the graphical notification area that identify current modes of operation of the speech recognizer.

21. The computer system of claim 20 further comprising a facility for displaying text in a window that identifies the current mode of operation of the speech recognizer.

22. The computer system of claim 21 wherein the facility for displaying text only displays in the window for a predetermined period of time.

23. A computer system comprising:

a video display for displaying video images;

an audio input device for receiving audio input from a user;

a speech recognizer for recognizing components of speech in audio input received by the audio input device; and

a first facility for displaying textual feedback to the user regarding processing of the received audio input by speech recognizer in a tool tip window that is only displayed for a predetermined period of time.

24. The computer system of claim 23 further comprising:

a second facility for providing a graphical notification area on the video display for displaying graphic notifications;

a third facility for displaying a graphical notification in the graphical notification area that provides feedback regarding processing of the received audio input by the speech recognizer.

25. In a computer system having a video display, an audio input device for receiving audio input from a user, a graphical user interface on the video display, said graphical user interface including a designated notification area for displaying graphical notifications and a speech recognizer for recognizing components of speech in audio input, said speech recognizer operating in a number of different modes of operation, a computer-readable medium holding computer-executable instructions for performing a method comprising the computer-implemented steps of:

determining a first current mode of operation of the speech recognizer;

displaying a first graphical notification of the first current mode of operation of the speech recognizer in the designated notification area;

determining a second current mode of operation of the speech recognizer; and

displaying a second graphical notification of the second current mode of the speech recognizer in the designated notification area, wherein the first notification differs from the second graphical notification.

26. The computer-readable medium of claim 25 wherein the first current mode of operation of the speech recognizer

is determined to be a listening mode in which the speech recognizer is listening for and processing audio input received at the audio input device.

27. The computer-readable medium of claim 25 wherein the first current mode of operation of the speech recognizer is determined to be a non-listening mode in which the speech recognizer is not listening for or processing the audio input received at the audio input device.

28. The computer-readable medium of claim 25 wherein the first current mode of operation of the speech recognizer is determined to be a standing-by mode in which the speech recognizer is listening for and processing only selected audio input received at the audio input device.

29. The computer-readable medium of claim 25 wherein the step of displaying the graphical notification of the first current mode of operation of the speech recognizer in the designated notification area comprises displaying an icon that identifies the first current mode of operation of the speech recognizer in the designated notification area.

30. The computer-readable medium of claim 25 wherein the computer system runs an operating system and the operating system provides the graphical user interface.

31. The computer-readable medium of claim 25 wherein the method further comprises the step of displaying text in a separate window that identifies the second current mode of operation.

32. The computer-readable medium of claim 31 wherein the method further comprises the step of terminating the display of the text and the separate window after expiration of a predetermined time period.

33. The computer-readable medium of claim 25 wherein the designated notification area does not obscure any other video output on the video display.

34. In a computer system having a video display, an audio input device for receiving audio input from a user and a speech recognizer for recognizing components of speech in audio input, a computer-readable medium holding computer-executable instructions for performing a method comprising the computer-implemented steps of:

receiving audio input from the user through the audio input device;

processing the received audio input with the speech recognizer to attempt to recognize components of speech in the received audio input; and

displaying textual feedback regarding the processing of the received audio input by the speech recognizer in a tool tip window that is only displayed for a predetermined period of time.

35. The computer-readable medium of claim 34 wherein the displayed textual feedback indicates that the received audio input was recognized by the speech recognizer.

36. The computer-readable medium of claim 35 wherein the displayed textual feedback indicates what the received audio input was recognized as.

37. The computer-readable medium of claim 34 wherein the displayed textual feedback indicates that the received audio input was not recognized by the speech recognizer.

38. The computer-readable medium of claim 34 wherein the displayed textual feedback indicates that there are problems with the received audio input that make it difficult for the speech recognizer to process the received audio input.

* * * * *

US006075534A

United States Patent [19]

VanBuskirk et al.

[11] Patent Number: 6,075,534

[45] Date of Patent: Jun. 13, 2000

[54] MULTIPLE FUNCTION GRAPHICAL USER INTERFACE MINIBAR FOR SPEECH RECOGNITION

[75] Inventors: **Ronald VanBuskirk**, Indiantown; **James R. Lewis**, Delray Beach, both of Fla.

[73] Assignee: **International Business Machines Corporation**, Armonk, N.Y.

[21] Appl. No.: **09/048,519**

[22] Filed: **Mar. 26, 1998**

[51] **Int. Cl.**[7] **G10L 9/06**; G06F 3/14

[52] **U.S. Cl.** **345/348**; 345/978; 704/275

[58] **Field of Search** 345/343, 326, 345/333, 334, 335, 336, 338, 339, 340, 347, 348, 349, 352, 978, 357; 704/235, 270, 275, 231

[56] **References Cited**

U.S. PATENT DOCUMENTS

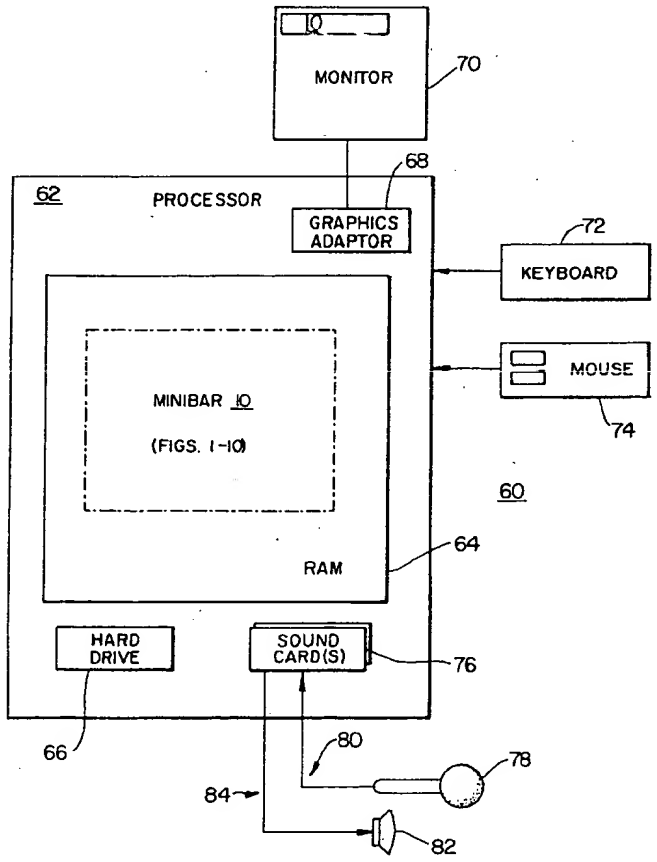
5,133,011	7/1992	McKiel, Jr.	704/276
5,386,494	1/1995	White	704/275
5,818,423	10/1998	Pugliese et al.	345/157
5,819,225	10/1998	Eastwood et al.	704/275
5,864,815	1/1999	Rozak et al.	704/275
5,864,819	1/1999	De Armas et al.	704/275
5,873,064	2/1999	De Armas et al.	704/275
5,890,122	3/1999	Van Kleeck et al.	704/275
5,903,870	5/1999	Kaufman	704/275
5,920,841	7/1999	Schottmuller et al.	704/275

Primary Examiner—Matthew M. Kim
Assistant Examiner—Crescelle N. dela Torre
Attorney, Agent, or Firm—Quarles & Brady LLP

[57] **ABSTRACT**

A multiple function graphical user interface for a speech recognition application adapted for generation by a computer programmed with a routine set of instructions, the interface comprising: an activatable icon defining an elongated screen display area, the icon having separately controllable foreground and background displays substantially coextensive with the display area; the background display being substantially fully changeable in response to user speech into a microphone having at least two states; text being displayable in the foreground display across substantially all of the display area; and, activation of the icon invoking a function related to the speech recognition application. A border can be selectively displayed around the elongated screen display area for indicating whether the speech application is in a navigation mode or a dictation mode.

31 Claims, 3 Drawing Sheets



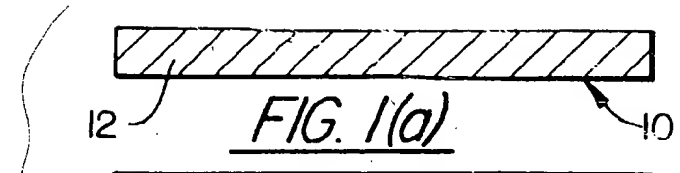


FIG. 1(a)

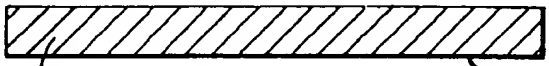


FIG. 1(b)

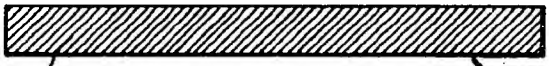


FIG. 1(c)

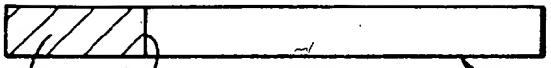


FIG. 2(a)

FIG. 2(b)

FIG. 3(a)

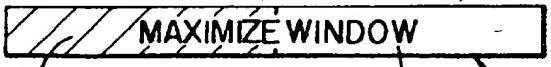


FIG. 3(b)

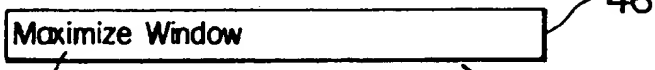


FIG. 4

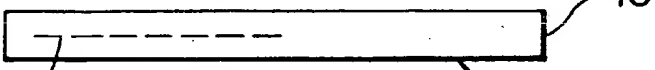


FIG. 5

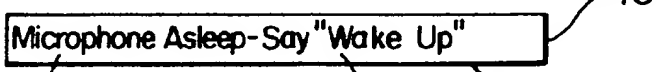


FIG. 6

FIG. 7

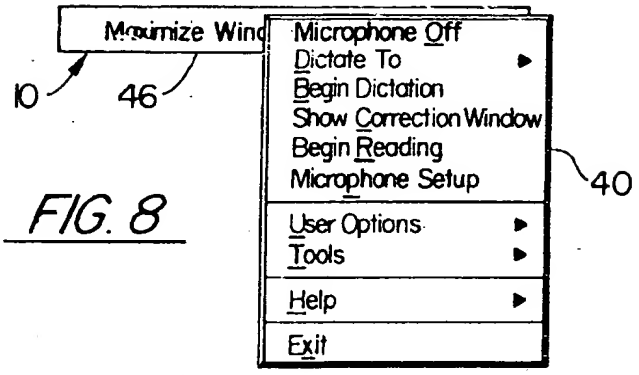


FIG. 8

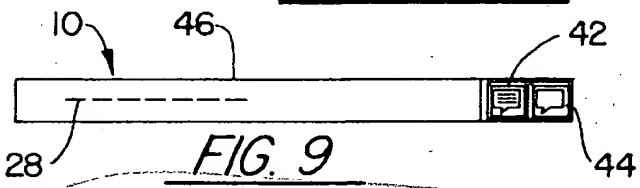


FIG. 9

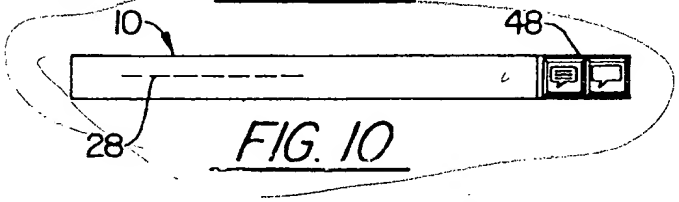


FIG. 10

FIG. 12(a)
(PRIOR ART)

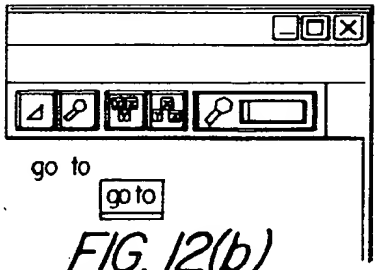


FIG. 12(b)
(PRIOR ART)

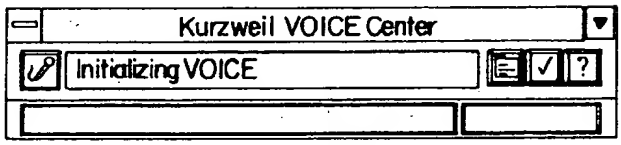


FIG. 13
(PRIOR ART)

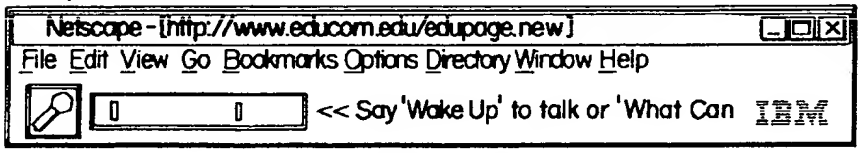


FIG. 14
(PRIOR ART)

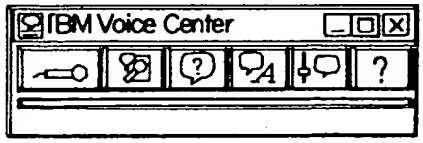


FIG. 15
(PRIOR ART)

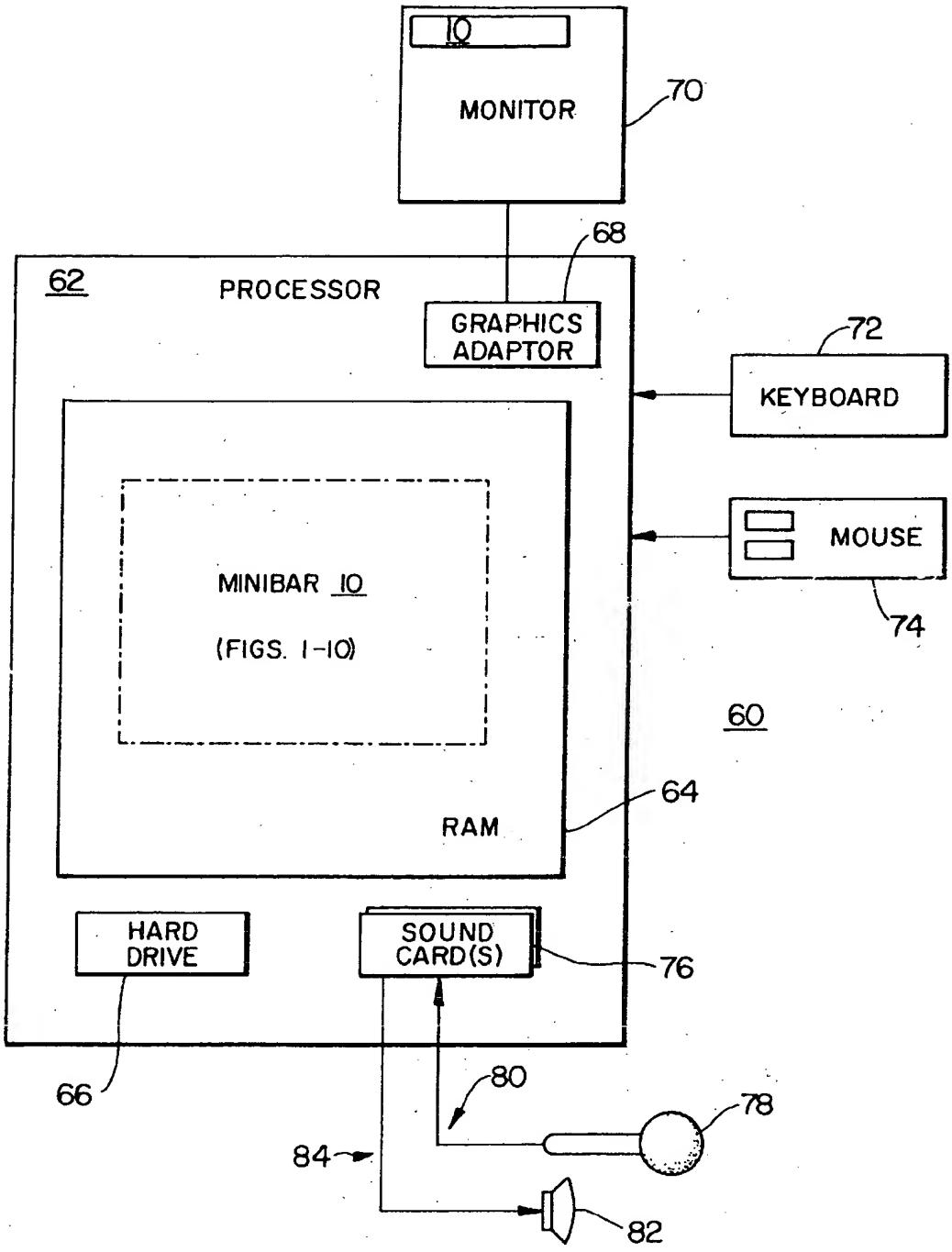

10
MONITOR
70
68
62
PROCESSOR
GRAPHICS ADAPTOR
72
KEYBOARD
MOUSE
74
MINIBAR 10
(FIGS. 1-10)
60
64
RAM
HARD DRIVE
SOUND CARD(S)
76
66
80
78
84
82

FIG. 11

MULTIPLE FUNCTION GRAPHICAL USER INTERFACE MINIBAR FOR SPEECH RECOGNITION

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates generally to the field of graphical user interfaces for speech applications, and in particular, to a multiple function graphical user interface for speech recognition which utilizes a minimum of screen space.

2. Description of Related Art

Presently, as developers add speech functions to computer graphical user interfaces, the speech function requires its own graphical interface. Speech users naturally want to minimize the amount of screen space occupied by the graphical user interface for speech recognition because the user can control the interface by voice. Generally speaking, the smaller the graphical user interface for speech recognition, the better.

Integrating speech recognition into a visually-based graphical user interface can be difficult. The difficulty stems from the fact that because the graphical user interface is visually-based, every encroachment upon this limited space reduces the free space for the user. However, the present level of speech recognition technology requires that the user be able to review several kinds of visual feedback. Therefore the challenge posed for developing an improved graphical speech interface is to present the minimal information required in the smallest space possible.

Current technology requires that the user have available the following information: 1) the state of the microphone speech system, that is on, off or asleep; 2) the last recognized phrase; 3) whether the application has speech focus; 4) feedback that the application is working; and, 5) status messages from the speech system. In addition the user must also have, at a minimum, the ability to control the microphone state with the mouse or other pointing tool, in case speech is not functioning. At best the user must have a mouse activatable equivalent for every speech command, because speech is not a deterministic input system.

Combining all of these functions into a truly minimal space is an elegant solution to a difficult problem. Various existing systems combine functionality in floating speech bars. Specific examples are instructive.

Existing systems combine their functionality in floating speech bars. The Dragon systems is shown in FIGS. **12**(*a*) and **12**(*b*). They combine the microphone and VU meter into one button, as shown in the upper right hand corner of FIG. **12**(*b*).

In VOICE from Kurzweil, as shown in FIG. **13**, all of the functions are combined on a tool bar and display the recognized text in a floating tool window.

The tool bars from VoiceType Connection and Voice Center, both available from IBM, are as shown in FIGS. **14** and **15** respectively.

None of the tool bar arrangements described above can truly be described as minimal, in terms of the necessary display area.

SUMMARY OF THE INVENTION

A minimal graphical user interface for speech recognition, also referred to herein as a minibar, in accordance with the inventive arrangements, combines the recognized text field, the on/off button, and the volume meter into one graphical user interface component.

A computer programmed with a routine set of instructions stored in a physical medium, for generating a multiple function graphical user interface for a speech recognition application in accordance with an inventive arrangement, comprises: means for displaying an activatable icon defining an elongated screen display area, the icon having separately controllable foreground and background displays substantially coextensive with the display area; means for changing substantially all of the background display in response to user speech into a microphone having at least two states; means for displaying text in the foreground display across substantially all of the display area; and, means responsive to activation of the icon for invoking a function related to the speech recognition application.

The changing means changes can advantageously change the background display by one of the following: changing from a first color to at least a second color responsive to variations of the volume of the user speech; generating a color ribbon having one fixed edge and one movable edge responsive to variations of the volume of the user speech; and, changing between lighter and darker shades of at least one color responsive to variations of the volume of the user speech.

The text displaying means, in conjunction with each embodiment of the changing means, can selectively display one or more of the following: each last recognized phrase of the user speech; status messages from the speech application; and, prompts for changing the states of the microphone.

In conjunction with each of the embodiments of the changing means, the function invoking means can display: a menu list of secondary speech functions responsive to activation of a pointing tool button when the icon is displayed; and, further icons for additional dictation functions can be displayed when dictation is active.

In conjunction with each of the embodiments noted above, the computer advantageously further comprises means for selectively generating a border around the elongated screen display area and around the further icons for indicating whether the speech application is in a navigation mode or a dictation mode.

A multiple function graphical user interface for a speech recognition application adapted for generation by a computer programmed with a routine set of instructions, in accordance with another inventive arrangement, comprises: an activatable icon defining an elongated screen display area, the icon having separately controllable foreground and background displays substantially coextensive with the display area; the background display being substantially fully changeable in response to user speech into a microphone having at least two states; text being displayable in the foreground display across substantially all of the display area; and, activation of the icon invoking a function related to the speech recognition application.

The background display can be advantageously changeable responsive to variations of the volume of the user speech in at least one of: from a first color to at least a second color; as a color ribbon having at least one movable edge; and, between lighter and darker shades of at least one color.

The text can advantageously includes at least one of: each last recognized phrase of the user speech; status messages from the speech application; and, prompts for changing the states of the microphone.

The invoked functions can advantageously include at least one of: displaying a menu list of secondary speech functions; and, displaying further icons for additional dictation functions.

The interface can advantageously further comprise a border selectively displayed around the elongated screen display area and around the further icons for indicating whether the speech application is in a navigation mode or a dictation mode.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. **1**(*a*), **1**(*b*) and **1**(*c*) show a first embodiment of a minimal graphical user interface tool bar in accordance with the inventive arrangements.

FIGS. **2**(*a*) and **2**(*b*) show a second embodiment of a minimal graphical user interface tool bar in accordance with the inventive arrangements.

FIGS. **3**(*a*) and **3**(*b*) show multiple functionality based on the first and second embodiments respectively.

FIG. **4** shows a graphical user interface tool bar, as in either of the first and second embodiments, in a resting state and displaying the last command.

FIG. **5** shows a graphical user interface tool bar, as in either of the first and second embodiments, in a resting state after a timeout function erases the last command shown in FIG. **4**.

FIG. **6** shows a graphical user interface tool bar, as in either of the first and second embodiments, in a resting state when the microphone is asleep.

FIG. **7** shows a graphical user interface tool bar, as in either of the first and second embodiments, in a resting state when the microphone is turned off.

FIG. **8** shows a graphical user interface tool bar, as in either of the first and second embodiments, with an appended pull-down menu.

FIG. **9** shows a graphical user interface tool bar, as in either of the first and second embodiments, in a navigation mode and with additional function buttons.

FIG. **10** shows a graphical user interface tool bar, as in either of the first and second embodiments, in a dictation mode and with additional function buttons.

FIG. **11** is a block diagram of a computer programmed with a routine set of instructions for generating a multiple function graphical user interface in accordance with the inventive arrangements, and as shown in FIGS. **1**–**10**.

FIGS. **12**(*a*), **12**(*b*) and **13**–**15** illustrate various speech tool bar arrangements known in the prior art.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The inventive arrangements embodied in a multiple function graphical user interface supply the required information in the smallest space possible. All of the following information can be provided.

The on, off or asleep state of the speech system is displayed. If the system is off, the preferred nature of the display is text, which tends to be more understandable than an icon.

The entire length of the minibar, which can change during the display and/or be user adjustable in length, can be devoted to displaying the last recognized phrase or command. Having extra room for the text field is particularly important for European, Middle Eastern and African languages.

Whether the application has speech focus is easily indicated with the large VU meter changing colors in the background, thus increasing the visibility of the active speech application. A large volume meter changing colors also alerts the user that the speech application is running.

Finally, being able to devote the entire length of the minibar to text also facilitates the display of status messages from the speech application.

A minimal graphical user interface for speech recognition combines the recognized text field, the on/off button, and the volume (VU) meter into one graphical user interface component. Part of the trick in generating a multiple function display in a single component advantageously relies on controlling the color or colors of the background of the component to represent the volume meter, which is difficult to show in black and white drawings.

FIG. **1**(*a*) shows a minibar **10** which is arbitrarily cross hatched to represent a light color **12**, for example yellow. FIG. **1**(*b*) shows the minibar **10** arbitrarily cross hatched with a denser pattern of lines to represent a darker color **14**, for example green. FIG. **1**(*c*) shows the minibar **10** arbitrarily cross hatched with a still denser pattern of lines to represent a still darker color **16**, for example red. The sequence of the yellow, green and red colors can be easily perceived as representing volume level of dictated speech. Alternatively, the progression of cross hatching density in FIGS. **1**(*a*), **1**(*b*) and **1**(*c*) can also represent lighter and darker shades of the same color, which can also be easily perceived as representing the volume level of dictated speech.

An alternative background color display is shown in FIGS. **2**(*a*) and **2**(*b*), in which the minibar **10** is arbitrarily cross hatched with a pattern of lines to illustrate displaying volume as a moving ribbon or thermometer **18**, wherein movement of the right edge **20** of the ribbon, for example, is easily perceived as representing the volume level of dictated speech.

The minibar **10** can display the last dictated phrase or command **24**, as shown in FIGS. **3**(*a*) and **3**(*b*), which differ from one another in the manner in which the background color display is implemented.

When no dictation has been received after a predetermined timeout or delay, the last phrase or command is erased. This condition can be displayed as shown in FIG. **5** by using a symbol **28** such as a dashed line.

When the microphone is asleep, and must be awakened for further dictation, the minibar **10** can appear as shown in FIG. **6**. A status message **30** indicates the microphone is asleep and a prompt **32** is the proper command to wake up the microphone.

When the microphone is turned off, and must be turned on for further dictation, the minibar **10** can appear as shown in FIG. **7**. A status message **34** indicates the microphone is turned off and a prompt **36** is the proper command to turn on the microphone.

The minibar **10** can be controlled, for example, by a mouse. Clicking with the left mouse button on the minibar can turn the minibar on and off. Clicking with the right mouse button can bring up a display of a list of secondary speech functions, for example a menu **40** as shown in FIG. **8**. When dictation is active in an application, additional dictation buttons **42** and **44**, for example for starting and stopping dictation respectively, will become visible as shown in FIGS. **9** and **10**. The minibar in FIG. **9** is in a navigation mode, indicated by a border **46**, preferably in a contrasting color, around the basic minibar. Navigation commands, also referred to as control commands, will appear in place of the dashed line **28** as they are recognized. The minibar in FIG. **10** is in a dictation mode, indicated by the border **48** around the start and stop dictation buttons **42** and **44**. The border **46** is also shown in FIGS. **4**, **5**, **6** and **8**,

as the dictation application is in the navigation mode. The borders, which can also be shown in different colors to represent status conditions, provide yet another function in a minimal screen area. A double left-click can be used to invoke a fast-path function, for example launching an audio adjustment application.

The minibar can be embedded into the Windows95 task bar; can be embedded into a tool bar in the application; can be embedded into the title bar of an application window; or, can be used as a floating window.

A computer system **60** is shown in block diagram form in FIG. **11**. The computer system is programmed with a set of instructions stored in a physical medium, for example a hard drive **66** and/or a random access memory (RAM) **64** of a central processor **62**, for generating the minibar **10** shown in FIGS. **1–10**. The computer has an audio interface **80**, **84** between one or more sound or audio cards **76** in the computer apparatus and each of a microphone **78** and a speaker **82**. Microphone **78** is intended to represent, for example, a headset microphone, a desktop microphone, a monitor microphone and a hand held microphone. Speaker **82** is intended to represent, for example, one or more internal speakers, one or more external speakers, one or more monitor speakers or a headset speaker. The computer system **60** has a graphics adapter **68** that generates the GUI displays explained above and shown in FIGS. **1–10**. The computer system further comprises a monitor **70**, a keyboard **72** and a mouse **74**. The dashed line box shown in random access memory **64** represents a programmed routine of instructions for generating the graphical user interface minibar, in accordance with the inventive arrangements. The instructions are stored in a physical medium embodied by hard drive **66** and loaded into another physical medium embodied by random access memory **64**. The programmed routine of instructions implements the steps of generating the minibar **10** as shown and described herein.

This method of displaying the basic speech information to the user minimizes the required screen space. The interface taught herein provides information a user requires as long as the chance of speech misrecognitions can occur, as long as there is a detectable delay in speech recognition, as long as the system has to provide the user with messages, and as long as the multiple windows paradigm is used on computers. This will continue to be the environment of speech applications for the foreseeable future.

What is claimed is:

1. A computer programmed with a routine set of instructions stored in a physical medium for generating a multiple function graphical user interface for a speech recognition application, comprising:

means for displaying an activatable icon having a border defining an elongated screen display area within said border, said icon having separately controllable foreground and background displays substantially coextensive with said display area, said foreground display superimposed on said background display;

means for changing substantially all of said background display in response to variations of the volume of user speech into a microphone having at least two states;

means for displaying text in said foreground display across substantially all of said display area; and,

means responsive to activation of said activatable icon for invoking a function related to said speech recognition application.

2. The computer of claim **1**, wherein said changing means changes said background display progressively from a first color to at least a second color responsive to variations of the volume of said user speech.

3. The computer of claim **1**, wherein said changing means changes said background by generating a color ribbon having one fixed edge and one movable edge responsive to variations of the volume of said user speech.

4. The computer of claim **1**, wherein said changing means changes said background display between lighter and darker shades of at least one color responsive to variations of the volume of said user speech.

5. The computer of claim **1**, wherein said text displaying means displays each last recognized phrase of said user speech.

6. The computer of claim **1**, wherein said text displaying means displays status messages from said speech application.

7. The computer of claim **1**, wherein said text displaying means displays prompts for changing said states of said microphone.

8. The computer of claim **1**, wherein said text displaying means selectively displays:

each last recognized phrase of said user speech;

status messages from said speech application; and,

prompts for changing said states of said microphone.

9. The computer of claim **2**, wherein said text displaying means selectively displays:

each last recognized phrase of said user speech;

status messages from said speech application; and,

prompts for changing said states of said microphone.

10. The computer of claim **3**, wherein said text displaying means selectively displays:

each last recognized phrase of said user speech;

status messages from said speech application; and,

prompts for changing said states of said microphone.

11. The computer of claim **4**, wherein said text displaying means selectively displays:

each last recognized phrase of said user speech;

status messages from said speech application; and,

prompts for changing said states of said microphone.

12. The computer of claim **1**, wherein said invoked function displays a menu list of secondary speech functions.

13. The computer of claim **12**, wherein said invoked function displays further icons for additional dictation functions.

14. The computer of claim **1**, wherein said invoked function displays further icons for additional dictation functions.

15. The computer of claim **13**, wherein said changing means changes said background display progressively from a first color to at least a second color responsive to variations of the volume of said user speech.

16. The computer of claim **13**, wherein said changing means changes said background by generating a color ribbon having one fixed edge and one movable edge responsive to variations of the volume of said user speech.

17. The computer of claim **13**, wherein said changing means changes said background display between lighter and darker shades of at least one color responsive to variations of the volume of said user speech.

18. The computer of claim **13**, wherein said text displaying means selectively displays:

each last recognized phrase of said user speech;

status messages from said speech application; and,

prompts for changing said states of said microphone.

19. The computer of claim 15, wherein said text displaying means selectively displays:

each last recognized phrase of said user speech;

status messages from said speech application; and,

prompts for changing said states of said microphone.

20. The computer of claim 16, wherein said text displaying means selectively displays:

each last recognized phrase of said user speech;

status messages from said speech application; and,

prompts for changing said states of said microphone.

21. The computer of claim 17, wherein said text displaying means selectively displays:

each last recognized phrase of said user speech;

status messages from said speech application; and,

prompts for changing said states of said microphone.

22. The computer of claim 1, further comprising means for selectively generating a border around said elongated screen display area for indicating whether said speech application is in a navigation mode or a dictation mode.

23. The computer of claim 8, further comprising means for selectively generating a border around said elongated screen display area for indicating whether said speech application is in a navigation mode or a dictation mode.

24. The computer of claim 13, further comprising means for selectively generating a border around said elongated screen display area and around said further icons for indicating whether said speech application is in a navigation mode or a dictation mode

25. The computer of claim 18, further comprising means for selectively generating a border around said elongated screen display area and around said further icons for indicating whether said speech application is in a navigation mode or a dictation mode.

26. A multiple function graphical user interface for a speech recognition application adapted for generation by a computer programmed with a routine set of instructions, said interface comprising:

an activatable icon having a border defining an elongated screen display area within said border, said icon having separately controllable foreground and background displays substantially coextensive with said display area, said foreground display superimposed on said background display;

said background display being substantially fully changeable in response to variations of the volume of user speech into a microphone having at least two states;

text being displayable in said foreground display across substantially all of said display area; and,

activation of said activatable icon invoking a function related to said speech recognition application.

27. The interface of claim 26, wherein said background display is changeable responsive to variations of the volume of said user speech in at least one of:

progressively from a first color to at least a second color;

as a color ribbon having at least one movable edge; and,

between lighter and darker shades of at least one color.

28. The interface of claim 26, wherein said text includes at least one of:

each last recognized phrase of said user speech;

status messages from said speech application; and,

prompts for changing said states of said microphone.

29. The interface of claim 26, wherein said invoked functions include at least one of:

displaying a menu list of secondary speech functions; and,

displaying further icons for additional dictation functions

30. The interface of claim 29, further comprising a border selectively displayed around said elongated screen display area and around said further icons for indicating whether said speech application is in a navigation mode or a dictation mode.

31. The interface of claim 26, further comprising a border selectively displayed around said elongated screen display area for indicating whether said speech application is in a navigation mode or a dictation mode.

* * * * *

US006233560B1

(12) **United States Patent**
Tannenbaum

(10) **Patent No.: US 6,233,560 B1**
(45) **Date of Patent: May 15, 2001**

(54) **METHOD AND APPARATUS FOR PRESENTING PROXIMAL FEEDBACK IN VOICE COMMAND SYSTEMS**

(75) Inventor: **Alan Richard Tannenbaum**, Austin, TX (US)

(73) Assignee: **International Business Machines Corporation**, Armonk, NY (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/213,857**

(22) Filed: **Dec. 16, 1998**

(51) **Int. Cl.**[7] **G10L 15/22**; G06F 3/16
(52) **U.S. Cl.** **704/275**; 345/347
(58) **Field of Search** 704/275; 345/347

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,451,895	*	5/1984	Sliwkowski	345/326
4,506,142	*	3/1985	Takano et al.	219/490
4,726,065		2/1988	Froessl .	
4,766,529		8/1988	Nakano et al. .	
4,783,759	*	11/1988	Borgendale et al.	345/347
5,027,406		6/1991	Roberts et al. .	
5,068,900		11/1991	Searcy et al. .	
5,133,011		7/1992	McKiel, Jr. .	
5,157,384		10/1992	Greanias et al. .	
5,222,146		6/1993	Bahl et al. .	
5,231,670		7/1993	Goldhor et al. .	
5,251,130	*	10/1993	Andrews et al.	704/3
5,305,244		4/1994	Newman et al. .	
5,386,494		1/1995	White .	
5,408,582		4/1995	Colier .	
5,428,707		6/1995	Gould et al. .	
5,452,397		9/1995	Ittycheriah et al. .	
5,465,317		11/1995	Epstein .	
5,500,920		3/1996	Kupiec .	
5,526,407		6/1996	Russell et al. .	
5,553,121		9/1996	Martin et al. .	
5,602,963		2/1997	Bissonnette et al. .	
5,604,840		2/1997	Asai et al. .	
5,632,002		5/1997	Hashimoto et al. .	
5,638,486		6/1997	Wang et al. .	
5,664,061		9/1997	Andreshak et al. .	
5,671,328		9/1997	Fitzpatrick et al. .	
5,698,834		12/1997	Worthington et al. .	
5,706,399		1/1998	Bareis .	
5,729,659		3/1998	Potter .	
5,864,815	*	1/1999	Rozak et al.	704/275
6,018,711	*	1/2000	French-St. George et al.	704/275

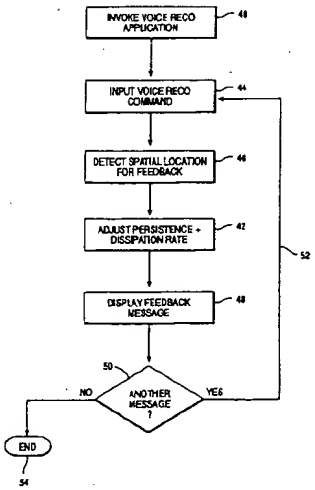
* cited by examiner

Primary Examiner—Tālivadis I. Šmits
(74) *Attorney, Agent, or Firm*—Robert M. Carwell

(57) **ABSTRACT**

In a voice actuated computer system voice command from an end user relevant to or promoted from a discrete location on a display screen are analyzed by the system. In response, a confirmation area is displayed on the screen at a location functionally related to the analyzed contents and context of the voice input or the screen location the utterance was prompted from. Within the confirmation area the computer interpretation of the utterance is displayed persisting and dissolving at selectively adjustable rates and times. Display of the recognized utterance is thereby placed in a confirmation area at variable locations where the user's focus is likely to be. Distractions are avoided associated with a fixed location confirmation area which obscures other content on the display screen and/or destroys end-user focus by requiring the eyes to shift from a location of current interest on the display screen to a different location wherein the confirmation is displayed. Persistence and dissolution time of the confirmation area and associated analyzed voice command displayed therein may vary automatically as a function of the degree of confidence of the voice recognition system that the command was accurately analyzed, and correctly recognized commands are automatically executed after display in the confirmation window.

23 Claims, 3 Drawing Sheets



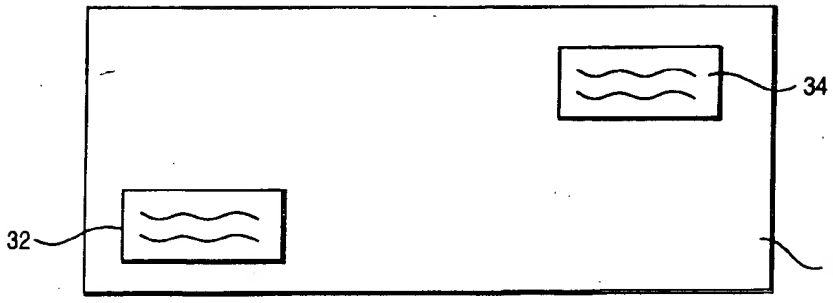


FIG. 1
PRIOR ART

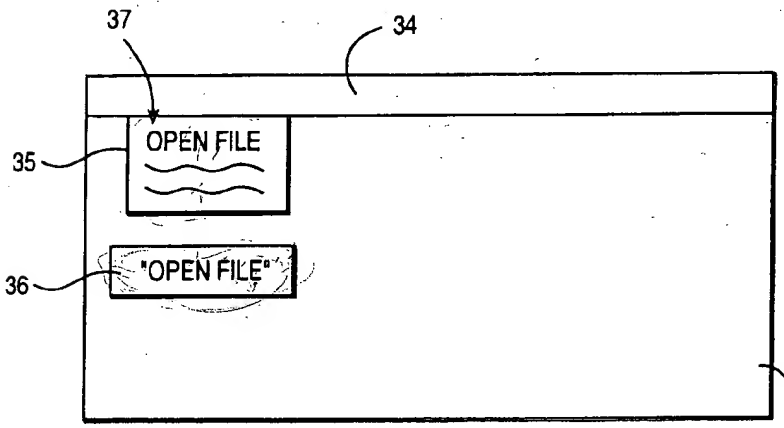


FIG. 2A

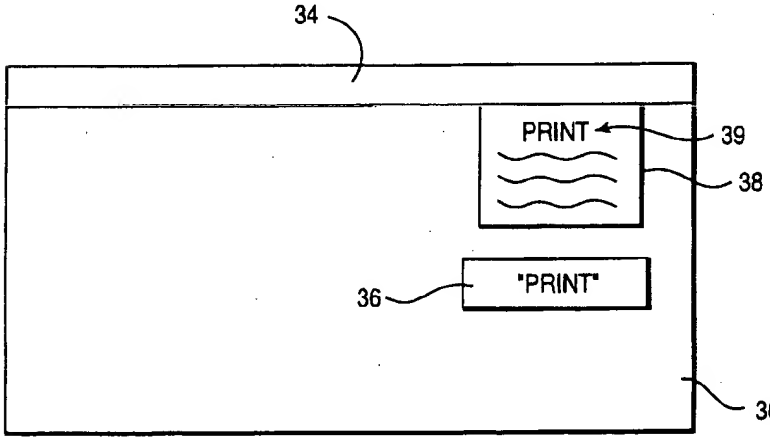


FIG. 2B

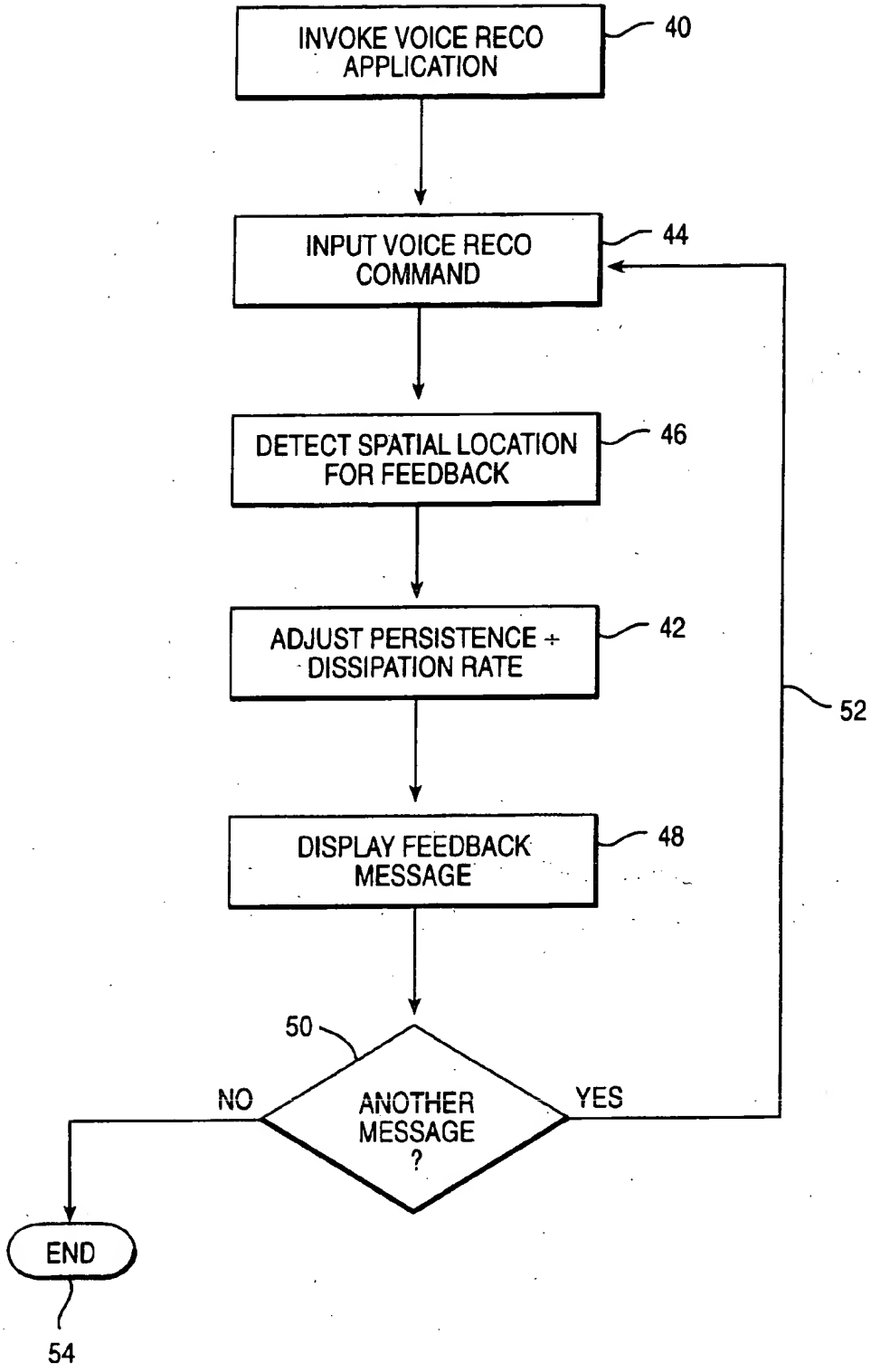

INVOKE VOICE RECO APPLICATION
40
INPUT VOICE RECO COMMAND
44
DETECT SPATIAL LOCATION FOR FEEDBACK
46
ADJUST PERSISTENCE + DISSIPATION RATE
42
52
DISPLAY FEEDBACK MESSAGE
48
50
NO
ANOTHER MESSAGE ?
YES
END
54

FIG. 3

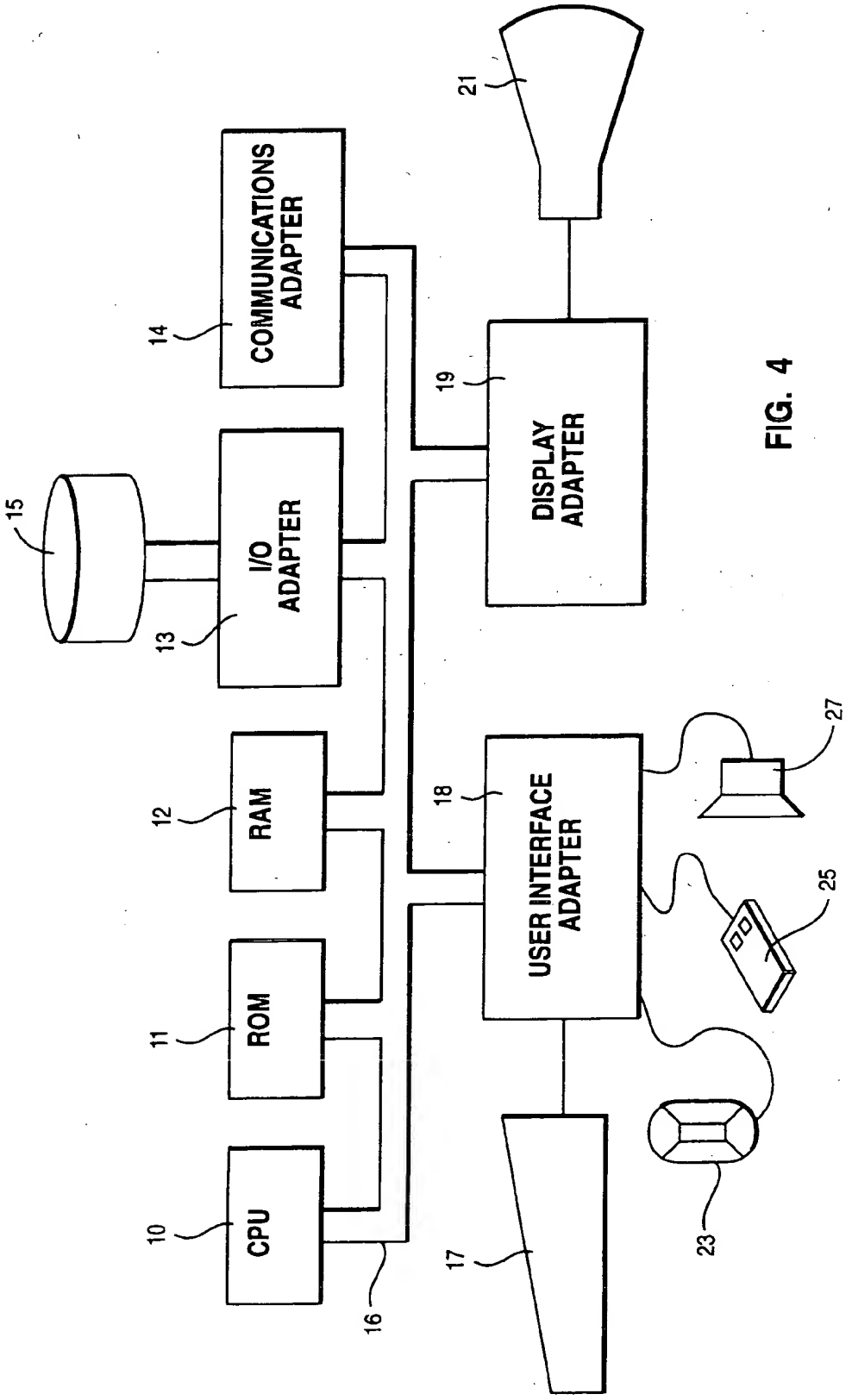

21
COMMUNICATIONS ADAPTER
14
19
DISPLAY ADAPTER
15
I/O ADAPTER
13
12
RAM
18
USER INTERFACE ADAPTER
27
25
11
ROM
23
10
CPU
17
16

FIG. 4

METHOD AND APPARATUS FOR PRESENTING PROXIMAL FEEDBACK IN VOICE COMMAND SYSTEMS

CROSS REFERENCE TO RELATED APPLICATIONS

The present application is related to the following applications: (1) application Ser. No. 09/213,856, filed Dec. 17, 1998, entitled "Speech Command Input Recognition System for Interactive Computer Display With Interpretation of Ancillary Relevant Speech Query, Terms Into Commands" (IBM Docket AT9-98-343); (2) application Ser. No. 09/213,858, filed Dec. 17, 1998, entitled "Speech Command Input Recognition System for Interactive Computer Display With Means for Concurrent and Modeless Distinguishing Between Speech Commands and Speech Queries for Locating Commands" (IBM Docket AT9-98-344); (3) application Ser. No. 09/213,846, filed Dec. 17, 1998, entitled "Speech Command Input Recognition System for Interactive Computer Display with Speech Controlled Display of Recognized Commands", (IBM Docket AT9-98-341); (4) application Ser. No. 09/213,845, filed Dec. 17, 1998, entitled "Speech Command Input Recognition System for Interactive Computer Controller Display with Speech Controller Display of Recognized Commands" (IBM Docket AT9-98-342). All of the above are assigned to a common assignee and filed on the same day as the present patent application.

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to voice command computer systems and, more particularly, to such systems which present command feedback to the end-user.

2. Background and Related Art

As the computer field has matured, vast improvements have been made in easing the manner in which end-users may interface with the systems. What was originally a crude command line interface requiring keyboard input of the end-user has now evolved to a rich panoply of techniques and devices for facilitating a more natural human interface with the computer. Representative examples of this include various pointing devices such as mice, track balls, touch-screens and voice activated interfaces made possible by advances in computerized speech recognition.

Such speech recognition falls into two major categories. First, systems have been developed for voice data input, examples of which include speech dictation or "speech-to-text" systems such as those marketed by the IBM Corporation and Dragon Systems. Yet a second category of computerized speech recognition technology is represented by command and control systems wherein end-user speech patterns are detected and analyzed. A specific form of this command and control technology is represented by a "desktop navigator", wherein a user may navigate through a computer system's graphical user interface (GUI) by merely speaking menu command words, labels on buttons, directional commands such as "up" and "left" to move the familiar mouse cursor, and the like.

Due to inherent voice recognition errors such speech recognition systems are prone to—partly because of the state of the technology and partly due to the variability of the user's speech patterns and memory of the correct inputs to utter, such systems typically reserve a dedicated, predetermined area of the GUI to display the system's interpretation of the user's utterances. The user, by inspecting this area, is looking for confirmation that the utterance he or she just made was in fact interpreted or acted upon correctly.

Current speech recognition systems display this "best guess" in a control window at a fixed position on the screen of the GUI in a control window. Although the spatial positioning of this control window may in some implementations be altered by the end-user as desired or is placed in a fixed location such as the title bar of the current application being executed, numerous problems nevertheless remain associated with this technology. These methods of displaying these interpreted commands attempt to place the confirmation area in a position so as to not block the visibility of other important portions of the application. However, in doing so, this may necessitate distracting and disrupting eye movement from the area of the screen where the end-user is focusing to the confirmation area of the voice recognition, whereupon the user must then recall and reposition his or her eyes at the area of the screen which had the user's focus prior to the interruption caused by display of the confirmation area. Moreover, current systems display the results of the speech perception system without taking into account interpretation of the command, as it is simply reported.

It will be readily appreciated that in sessions of any length with such a voice recognition computer system these drawbacks can become extremely tiresome both physically and mentally and can severely impact the productivity of the end-user in a manner as to almost effectively negate the aforementioned benefits to voice recognition or navigation systems. Current methods and technology have simply been unable to eliminate these serious problems of maintaining visibility of important areas of the display and the focusing and refocusing just described.

Accordingly, such a voice command and control system was highly desired which could avoid the distractions to end-users and the degradation of visibility of applications presented on GUI user interfaces.

SUMMARY OF THE INVENTION

In a voice actuated computer system, voice input from an end-user, such as command and/or control utterances relevant to or prompted from a discrete location on a display screen, is analyzed by the system. In a response to such analysis, a confirmation area is thereby displayed on the display screen in a location functionally related to the analyzed contents of the voice input or the location the utterance was prompted from. Within the confirmation area the computer system's interpretation of the utterance is displayed for a preselected time and then gradually dissolved also at a preselected rate, such persistence and dissolution being selectively adjustable by the end-user. In this manner, display of the interpretation of the command or control utterance will be placed in a confirmation area which is in turn spatially positioned on the display screen at variable locations functionally related to the content of the command or control utterance so as to be proximal to the location on the display screen which may have had the user's eye focus and which gave rise to the utterance. In this manner, distractions are thereby avoided associated with a fixed location confirmation in the prior art (1) which obscures content of interest on the display screen and/or (2) destroys end-user focus by requiring the eyes to shift from positions of interest on the display screen to a different location wherein the confirmation is displayed. An important factor in the invention is the interpretation of the utterance and therefore the true target of the command. Feedback will be preferably withheld, momentarily in most cases, until

such correct interpretation to application functionality can be determined.

In a more general case, the invention is not intended to be limited to the prior art programmer's graphical model comprised of interaction and association with cursor targets, menu items, pointing devices, and toolbars. Rather, the invention contemplates a visual mechanism for display of recognized speech commands in the form of natural language and resulting feedback actions proximal to the area of focus in a graphical user interface which gave rise to and prompted the speech command. Such feedback actions may include text over text, visuals, color change, animation, and gradual reduction of an image or message fading into the target area automatically so as to provide feedback to the end-user that the correct spoken action has been taken in the desired focus or target area.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** is a representative display screen in a voice activated computer system demonstrating the method of the prior art for displaying confirmation areas.

FIGS. **2A** and **2B** are illustrations of display screens in accordance with the invention for improved displays of confirmation areas associated with voice command and control systems.

FIG. **3** is a flow diagram implementable in program code executing in the computer system of the invention for effecting the improved confirmation area depicted in FIG. **2**.

FIG. **4** is a simplified block diagram depicting a voice actuated computer system in which the features of the invention are advantageously implemented.

DETAILED DESCRIPTION

Turning first to FIG. **1**, depicted therein is a simplified illustration of a representative display **30** which might occur on the monitor **39** of the system of FIG. **4** in accordance with the prior art. In this system, a user might be interacting with and accordingly have eyes focused upon an area **32** of the screen **30** which may prompt voice command to be analyzed and acted upon in the system of FIG. **4**. Once the voice recognition system makes an attempt to recognize such an utterance, it has been conventional to display the computer's interpretation or translation of this command in a confirmation area **34**. For example, if the user is viewing an area **32** of an application in which there is menu of voice activatable choices such as "open dictionary", he or she may consequently utter this command. In response thereto, the computer system, after recognizing the command or its best guess as to an interpretation, this interpretation (such as "open dictionary" if the utterance was correctly so-interpreted) would appear in this confirmation area **34**.

A significant deficiency in this operation of voice command systems previously alluded to may be seen clearly depicted with reference to this prior art screen **30** of FIG. **1**. It will be noted that whereas the end-user's focus was on area **32**, in response to a voice command prompted thereby he or she will receive feedback as to how the command has been interpreted in this confirmation area **34** which is spatially separated a significant distance on the user interface screen from the area **32** which prompted the command in the first place. This causes the user to have to refocus his or her eyes at the diagonally opposite end of the screen **30** at the confirmation area **34** in order to discern whether or not the voice command was correctly interpreted. If so, the user must then cause his or her eyes to traverse back to area **32** and refocus and to regain concentration on the content of this area **32**. One reason historically for locating this confirmation area **34** which typically persists for a great deal of time in a fixed and out of the way location on the screen **30** is so that does not thereby obscure other relevant portions of the screen **30** wherein other informational content of the application is to be displayed.

Turning to FIGS. **2A** and **2B**, displayed therein are correlative examples of screens **30** corresponding to that of the prior art screen **30** of FIG. **1**, wherein the features and benefits of the invention may be seen clearly depicted therein. Referring first to FIG. **2A**, again a screen **30** on which is displayed a representative screen from a voice actuated application executing in the computer system of FIG. **4** is shown. In this screen **30** it will be noted that a conventional tool bar **34** is shown which may have a pulldown menu **35** associated therewith in the upper left-hand corner of the screen **30**. One voice recognizable command shown in this pulldown menu **35** might be "open file" **37**. When the end-user utters this "open file" command **37**, as in the case of the interface of FIG. **1**, the familiar confirmation area **36** may be made to appear proximal to the menu **35** wherein the voice-recognized utterance "open file" will appear (or whatever the system recognizes) thereby providing feedback to the user as to whether the user's speech was correctly recognized.

Similarly, turning to FIG. **2B**, yet another pulldown menu **38** may appear associated with the task or tool bar **34** at the upper portion of the screen **30** associated with yet another series of voice recognizable commands. It will be noted that this pulldown menu is at a location on the screen **30** different from that of FIG. **2A**, and that this pulldown menu might have a different voice actuatable command such as "print" **39** displayed in this pulldown menu. Upon the user uttering this voice actuatable command "print", in like manner to the case with respect to FIGS. **1** and **2A**, in this FIG. **2B** will be seen that a confirmation area **36** appears on the display screen **30** proximal to where the voice actuatable "print" **39** command appears.

A comparison of these FIGS. **2A** and **2B** with the prior art user interface of FIG. **1** reveals an important and significant difference. As previously described, in the case of the prior art FIG. **1** interface, the area of concentration of the end-user **32** may be spatially a significant distance from where the default confirmation area **34** appears, giving rise to the associated undesirable results of disrupting concentration, causing the necessity of eye refocus and the like as the eye traverses between areas **32** and **34**.

In contrast, however, it will be noted that in the user interface examples of FIGS. **2A** and **2B**, a significant difference is that this confirmation area **36** is displayed on the screen **30** in an area proximal to the area on the screen which gave rise to the voice command which in turn caused display of the confirmation area **36** and wherein the user's eyes are thus most likely to be focused. Thus, for example, in FIG. **2A**, the confirmation area **36** is proximal, e.g., just below the location on screen **30** where the voice actuatable "open file" command **35** is displayed in the pulldown menu which, when uttered, gave rise to display of this confirmation area **36** on the screen and the recognized "open file".. In like manner, in the illustration of FIG. **2B**, this confirmation area **36** now appears at a different location on the screen from that of FIG. **2A**, namely at a location proximal to the "print" voice actuatable command **39** associated with pulldown menu **38** which itself appears in a different location from the pulldown menu **35** including the "file" command of FIG. **2A**. In this manner, the user's eyes do not have to traverse

a disrupting significant distance from the location on the display screen which prompted the voice actuatable command in order to verify that the command was correctly interpreted as viewed in the confirmation area **36**.

It is yet a further significant feature of the invention that in the course of the voice navigation program which analyzes and displays the interpreted command in the confirmation area **36** that a context-sensitive analysis function which may include elements of artificial intelligence as desired and appropriate, may be built into the voice recognition system so as to position this confirmation area **36** proximal on the screen **30** to the intended and appropriate target of the command also displayed on the screen **30** which gave rise to display of the confirmation area and the interpreted command displayed therein. Thus, for example, upon the end-user uttering "open file", this subroutine or program feature in the voice recognition program executing on the system of FIG. **4** will determine that because this "open file" command is currently being displayed in the upper left-hand corner of the screen **30** that the confirmation area **36** should therefore appropriately be displayed somewhere proximal to where this "open file" command appears in the drop-down menu **35** of the toolbar **34**. Similarly, upon the voice recognition system detecting that the command "print" **39** has been uttered and correlating this to the fact that this command is displayed in the upper right corner of the display **30**, the system of the invention will thereby determine that, in contrast to the location of the display of the confirmation area **36** of FIG. **2A**, it would now be more appropriate to display this confirmation area **36** in FIG. **2B** at a different location on the display screen, e.g., at a location proximal to where this "print" command **39** appeared which gave rise to the "print" utterance from the end-user.

Turning now to FIG. **3**, depicted therein is a flow diagram illustrating how program code would be provided executing on the system of FIG. **4** to implement the aforementioned features of the invention. First, it will be assumed that the end-user has invoked a voice recognition or navigation application, as shown at box **40** which will be executing on the system of FIG. **4**. It may be desirable to the end-user to adjust the persistence and dissipation of the confirmation area **36** prior to continuing with the process of FIG. **3**. Accordingly, as shown by the dotted lines and the box **42**, this program code may be adapted to accommodate this feature. Alternatively, it will be appreciated that a feature of the invention may be to selectively alter this persistence and dissipation automatically, for example, as a function of the degree of confidence with which the voice recognition system has recognized the utterance of the end-user. If, for example, the voice recognition system has almost a 100% confidence factor that the words "close application" have been recognized, the persistence time and dissipation time of the confirmation message would desirably be lessened substantially in that it is less likely that this feedback would be of significance to the end-user. On the other hand, if the voice recognition application had difficulty in recognizing an utterance, it would be desirable to automatically lengthen the persistence and dissipation time of the confirmation message. This will allow the end-user more time to notice that the command has not been interpreted correctly and to take corrective actions such as manually entering the correct desired command in the confirmation area or by manually effecting the command through use of the keyboard or pointing device, or by invoking an undo feature.

Continuing with FIG. **3**, once the voice recognition application program has been invoked and is executing, it will be assumed that the end-user will input a voice recognition command shown at box **44** in response to viewing a correlative command displayed on the display **30**. Such input may be made by means of a microphone **28** shown in the system of FIG. **4** and in response to reviewing one or more display screens **30** occurring on the monitor **39** of FIG. **4**.

Once this voice recognition command has been uttered, the system of FIG. **4** will thereafter detect an appropriate spatial location for providing feedback to the end-user in a confirmation area **36**, this step being illustrated by block **46** of FIG. **3**. It will be recalled that this implementation of this functional block may include, as previously described, various factors as appropriate—including elements of artificial intelligence which have been sensing prior user interaction with the application. Additionally, such detection at block **46** may further include analysis of which voice actuatable commands are currently being displayed on the display screen **30**, and may also incorporate intelligent predictions as to where the end-user's eye and intellectual focus may next occur on the display screen **30** based upon analysis of prior interaction with the voice actuatable program.

Continuing with FIG. **3**, once the program code has detected the desirable spatial location for feedback in the confirmation area **36**, the system will thereafter cause display of an appropriate feedback message **48** in a confirmation area **36** located on the screen **30** (FIGS. **2A–2B**) based upon the analysis which previously transpired with reference to the function of box **46**. It will be noted that one appropriate such feedback message might be simply displaying the alpha numeric version of the uttered command or instruction from the end-user as interpreted by the voice recognition program, e.g., displaying the words "open file" in the confirmation area **36** upon detecting that it appears the end-user has uttered the words "open file" in response to viewing this as a voice actuatable command choice in the pop-up menu **35**.

Once this "best guess" interpretation of the uttered voice command has thereby been displayed, **48**, the process of FIG. **3** thereafter queries whether another message or command has been uttered by the end-user, shown by the decision block **50** of FIG. **3**. If another message has been detected, flow exits to the left of block **50** and is routed along path **52** back to block **44** wherein this next voice command utterance input will be analyzed. Flow then continues vertically downward along the flow diagram of FIG. **3**, whereupon this next desired spatial location for the confirmation area **36** for this next utterance is detected and the particular command analyzed. If, on the other hand, in response to the query at decision block **50**, it has been determined that another voice command has not been entered or will not be entered (for example, because the voice navigable application has been closed), the process exits to the right of decision block **50** and the process ends, **54**.

Before describing a block diagram of a computer system in which the invention may be advantageously employed, a few additional points must be noted. It should be readily apparent from the foregoing that once the confirmation area **36** has been extinguished after an appropriate length of persistence and dissipation, if the command has been correctly recognized it will automatically be executed by the system. If, on the other hand, it has not been correctly recognized, while the confirmation area still persists and before it has dissipated, the end-user may (depending upon a user-determined preference) override the recognized command (or cause execution of a correct command in the event the command has not been recognized) by means of manual entry in the keyboard, or use of a pointing device as required, or the utterance of a special voice command such as "STOP".

Yet an additional point is that although in the foregoing illustrations, voice recognition has been employed with respect to predetermined command choices for the end-user to select from a displayed pulldown menu or the like, the invention is not intended to be so limited. Accordingly, there may be voice actuatable actions not associated with messages appearing on the screen. Representative examples of this might be directional controls such as "move cursor down" or "enlarge figure". Such uttered directions may also, in like manner to a menu of command choices displayed on the screen, be recognizable by the voice system and acted upon, also with a confirmation of the words uttered appearing in a confirmation area **36** prior to execution of the recognized command or correction thereof to the keyboard or pointing device. Also, in keeping with a fundamental concept of the invention, even in such cases wherein a menu of spoken commands to select from is not present, it is contemplated that the voice recognition system will be command-context sensitive in the sense that it may intelligently determine where to place the confirmation area **36** dependent upon the words uttered by the end-user. For example, if the end-user had uttered "move cursor to the right two inches", the system could determine by recognizing the word "cursor" that some action associated with the cursor was desired—such as a move, shape change, or the like. Accordingly, the system would automatically display the confirmation area **36** with the recognized command displayed proximal to the current location of the cursor on the display screen **30**. This, of course, is in recognition of the likelihood that the end-user will have focused his or her eyes immediately preceding the command on the cursor and thereby may be expecting a confirmation message to appear somewhere proximal to the current cursor location.

FIG. **4** illustrates a preferred embodiment of a computer system which may advantageously employ the improved pointing device resolution system of the present invention. The system comprises a CPU **10**, read only memory (ROM) **11**, random access memory (RAM) **12**, I/O adapter **13**, user interface adapter **18**, communications adapter **14**, and display adapter **19**, all interconnected via a common address/data and control path or bus **16**. Each of the above components accesses the common bus utilizing conventional techniques known to those of ordinary skill in the art, and includes such methods as dedicating particular address ranges to each component in the system, with the CPU being the busmaster. As is further shown in FIG. **4**, these external devices such as DASD **15** interface to a common bus **16** through respective adapters such as I/O adapter **13**. Other external devices, such as the display **21**, similarly use their respective adapter such as display adapter **19** to provide data flow between the bus **16** and the display **21** or other device. Various user interface means are provided for interconnection and use with the user interface adapter **18**, which, in the figure has attached thereto representative user input devices such as joy stick **23**, mouse **25**, keyboard **17**, and speaker and/or microphone **27**. Each of these units is well known in as such and accordingly will not be described herein.

The invention admits to implementation on essentially any computer system and corresponding microprocessor, such as the RS/6000™, RISC-based workstations and personal computers of the IBM Corporation executing the AIX™ and OS/2™ operating systems, respectively, or similar machines of other vendors, which include for example in the case of an RS/6000 workstation a 604 PowerPC™ RISC chip. (RS/6000, IBM, AIX, OS/2 and PowerPC are trademarks of the IBM Corporation).

Contained with the CPU **10** of FIG. **4**, typically is one or more microprocessors which performs the system address, data, and control processing functions necessary for correct operation of the system of FIG. **4**. Although the invention admits to application to various microprocessor designs, in the embodiment disclosed herein, the microprocessor takes the form of a PowerPC 604 microprocessor manufactured by the IBM Corporation, which is a species of microprocessor known as a reduced instruction set computer (RISC) microprocessor. Further details regarding the architecture and operation of such a microprocessor may be obtained from the PowerPC 604 RISC Microprocessor Users Manual, Document #MPC604UM/AD, November, 1994, copyright IBM Corporation, which is incorporated herein by reference.

In the context of the invention, the user will view various objects such as a cursor and pop up or pop down menus on the display **21** which may be manipulated by means of various pointing devices such as the mouse **25** and voice activated navigation. Program code associated with the user interface adapter **18** by way of a device driver for the pointing device **25** and microphone **27** in conjunction with operating environment and application code resident in RAM **12** and/or DASD **15** will facilitate and enable movement of a cursor on the display screen **21** responsive to and in association with correlative voice commands spoken into microphone **27**.

It will be understood from the foregoing description that various modifications and changes may be made in the preferred embodiment of the present invention without departing from its true spirit. It is intended that this description is for purposes of illustration only and should not be construed in a limiting sense. The scope of this invention should be limited only by the language of the following claims.

What is claimed is:

1. A method for presenting feedback on a display in a voice command recognition computer system executing a voice recognition application, comprising:

recognizing a speech command;

selecting one of a plurality of positions on said display as a function of said recognized speech command;

displaying feedback at a confirmation area proximal to said one of said positions and corresponding to said command;

extinguishing said feedback at a time after said displaying;

automatically executing said command after said displaying; and

wherein timing of said extinguishing is dynamically selected and wherein said feedback is selected from a group comprising a color change, animation, message, or reducing image on said display.

2. The method of claim **1** further including

determining a confidence factor corresponding to accuracy of said recognizing; and

wherein said time of extinguishing corresponds to said confidence factor.

3. The method of claim **2** including

depicting on said display information corresponding to said speech command at a location on said display;

wherein said speech command corresponds to said information; and

wherein said one of said positions corresponds to said location.

4. The method of claim **2** including

analyzing said speech command; and

wherein said one of said positions is a function of said analyzing.

5. The method of claim 2 wherein said computer system includes

a graphical user interface on said display and a pointing device; and

wherein said speech command corresponds to a command displayed on said graphical user interface actuatable by said pointing device.

6. The method of claim 2 wherein said confidence factor and said timing of extinguishing are inversely related.

7. The method of claim 1 further including

recognizing a next speech command;

selecting a next one of a plurality of positions on said display screen as a function of said recognized next speech command; and

displaying next feedback at a next confirmation area proximal to said next one of said positions and corresponding to said next speech command.

8. An apparatus for presenting feedback on a display in a voice command recognition computer system executing a voice recognition application, comprising:

means for recognizing a speech command;

means for selecting one of a plurality of positions on said display as a function of said recognized speech command;

means for displaying feedback at a confirmation area proximal to said one of said positions and corresponding to said command;

means for extinguishing said feedback at a time after said displaying; means for automatically executing said command after said displaying; and

wherein said timing of said extinguishing is dynamically selected and wherein said feedback is selected from a group comprising a color change, animation, message, or reducing image on said display.

9. The apparatus of claim 8 further including

means for determining a confidence factor corresponding to accuracy of said detecting; and

wherein said time of extinguishing corresponds to said confidence factor.

10. The apparatus of claim 9 including

means for depicting on said display information corresponding to said speech command at a location on said display;

wherein said speech command corresponds to said information; and

wherein said one of said positions corresponds to said location.

11. The apparatus of claim 9 including

means for analyzing said speech command; and

wherein said one of said positions is a function of said analyzing.

12. The apparatus of claim 9 wherein said computer system includes

a graphical user interface on said display and a pointing device; and

wherein said speech command corresponds to a command displayed on said graphical user interface actuatable by said pointing device.

13. The apparatus of claim 9 wherein said computer system includes

a graphical user interface on said display and a pointing device; and

wherein said speech command corresponds to a non-visible command interpreted by an application running on said graphical environment.

14. The apparatus of claim 9 wherein said confidence factor and said timing of extinguishing are inversely related; and wherein said computer system further includes

a graphical user interface on said display and a pointing device; and

wherein said speech command corresponds to a non-visible command interpreted by an application executing in said graphical user interface.

15. The apparatus of claim 8 further including

means for recognizing a next speech command;

means for selecting a next one of a plurality of positions on said display screen as a function of said recognized next speech command; and

means for displaying next feedback at a next confirmation area proximal to said next one of said positions and corresponding to said next speech command.

16. A program product for presenting feedback on a display in a voice command recognition computer system executing a voice recognition application, comprising:

program code means for recognizing a speech command;

program code means for selecting one of a plurality of positions on said display as a function of said recognized speech command; and

program code means for displaying feedback at a confirmation area proximal to said one of said positions and corresponding to said command;

program code means for extinguishing said feedback at a time after said displaying;

program code means for automatically executing said command after said displaying; and

wherein said timing of said extinguishing is dynamically selected and wherein said feedback is selected from a group comprising a color change, animation, message, or reducing image on said display.

17. The program product of claim 16 further including

program code means for determining a confidence factor corresponding to accuracy of said recognizing a speech command; and

wherein said timing of extinguishing corresponds to said confidence factor.

18. The program product of claim 17 including

program code means for depicting on said display information corresponding to said speech command at a location on said display;

wherein said speech command corresponds to said information; and

wherein said one of said positions corresponds to said location.

19. The program product of claim 17 including

program code means for analyzing said speech command; and

wherein said one of said positions is a function of said analyzing.

20. The program product of claim 17 wherein said computer system includes

a graphical user interface on said display and a pointing device; and

wherein said speech command corresponds to a command displayed on said graphical user interface actuatable by said pointing device.

21. The program product of claim 17 wherein said computer system includes

a graphical user interface on said display and a pointing device; and

wherein said speech command corresponds to a non-visible command interpreted by an application running on said graphical environment.

22. The program product of claim 17 wherein said confidence factor and said timing of extinguishing are inversely related.

23. The program product of claim 16 further including

program code means for recognizing a next speech command;

program code means for selecting a next one of a plurality of positions on said display screen as a function of said recognized next speech command; and

program code means for displaying next feedback at a next confirmation area proximal to said next one of said positions and corresponding to said next speech command.

* * * * *

US006018711A

United States Patent [19]

French-St. George et al.

[11] Patent Number: 6,018,711

[45] Date of Patent: Jan. 25, 2000

[54] **COMMUNICATION SYSTEM USER INTERFACE WITH ANIMATED REPRESENTATION OF TIME REMAINING FOR INPUT TO RECOGNIZER**

[75] Inventors: **Marilyn French-St. George**, Alcove; **Frederic Fernand Trasmundi**, Ottawa; **Jobe L.W. Roberts**, Wakefield, all of Canada

[73] Assignee: **Nortel Networks Corporation**, Montreal, Canada

[21] Appl. No.: **09/063,007**

[22] Filed: **Apr. 21, 1998**

[51] **Int. Cl.**[7] **G10L 5/00**

[52] **U.S. Cl.** **704/275**; 704/270; 455/231

[58] **Field of Search** 704/270, 275; 455/231

[56] **References Cited**

PUBLICATIONS

Jian Wang, "Integration of eye-gaze, voice and manual response in multimodal user interface," Proc. IEEE International Conference on Systems, Man, and Cybernetics, vol. 5, p. 3938–3942, Oct. 1995.

M. Billinghurst and J. Savage, "Adding intelligence to the interface," Proc. IEEE International Symposium on Virtual Reality, p. 168–175, Mar. 1996.

Rajeev Sharma, Vladimir I. Pavlovic, and Thomas S. Huang, "Toward multimodal human–computer interface,"Proc. IEEE, vol. 86, No. 5, p. 853–869, May 1998.

Ben Salem, Rob Yates, and Reza Saatchi, "Current trends in multimodal input recognition," IEE Colloquium on Virtual Reality Personal Mobile and Practical Applications (Ref. No. 1998/454), p. 3/1–3/6, Oct. 1998.

Primary Examiner—David R. Hudspeth

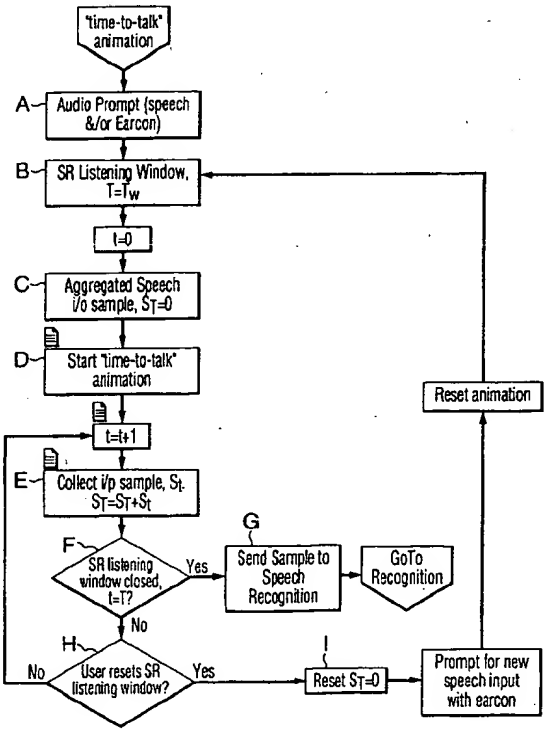
Assistant Examiner—Tālivaldis Ivars Šmits

Attorney, Agent, or Firm—Angela C. de Wilton

[57] **ABSTRACT**

A system and method for management of an interface for communications systems and devices using a recognizer having a limited time duration recognition window, based on mapping available time to a spatial representation. The system has particular application to management of a speech interface for a speech recognizer with a limited time duration recognition window. When the speech recognizer is on and a recognition window is opened, a timing signal is sent to the device and processed to provide spatial representation of the time remaining, e.g. in the form of an animation on a graphical user interface, or a recognizable pattern of stimulation using a haptic interface. As the recognition window advances and closes, the animation or pattern also changes and closes, e.g. diminishes in size, to provide the user with spatial information indicative of the time left in the recognition window. A reset feature allows users to reopen the recognition window to delete and correct input, while the window is still open, or after closing the window, and before the system advances to an erroneous state. The system is particularly applicable to speech recognizers having a limited recognition window that access speech recognition engines resident on a network or at a terminal device.

29 Claims, 4 Drawing Sheets



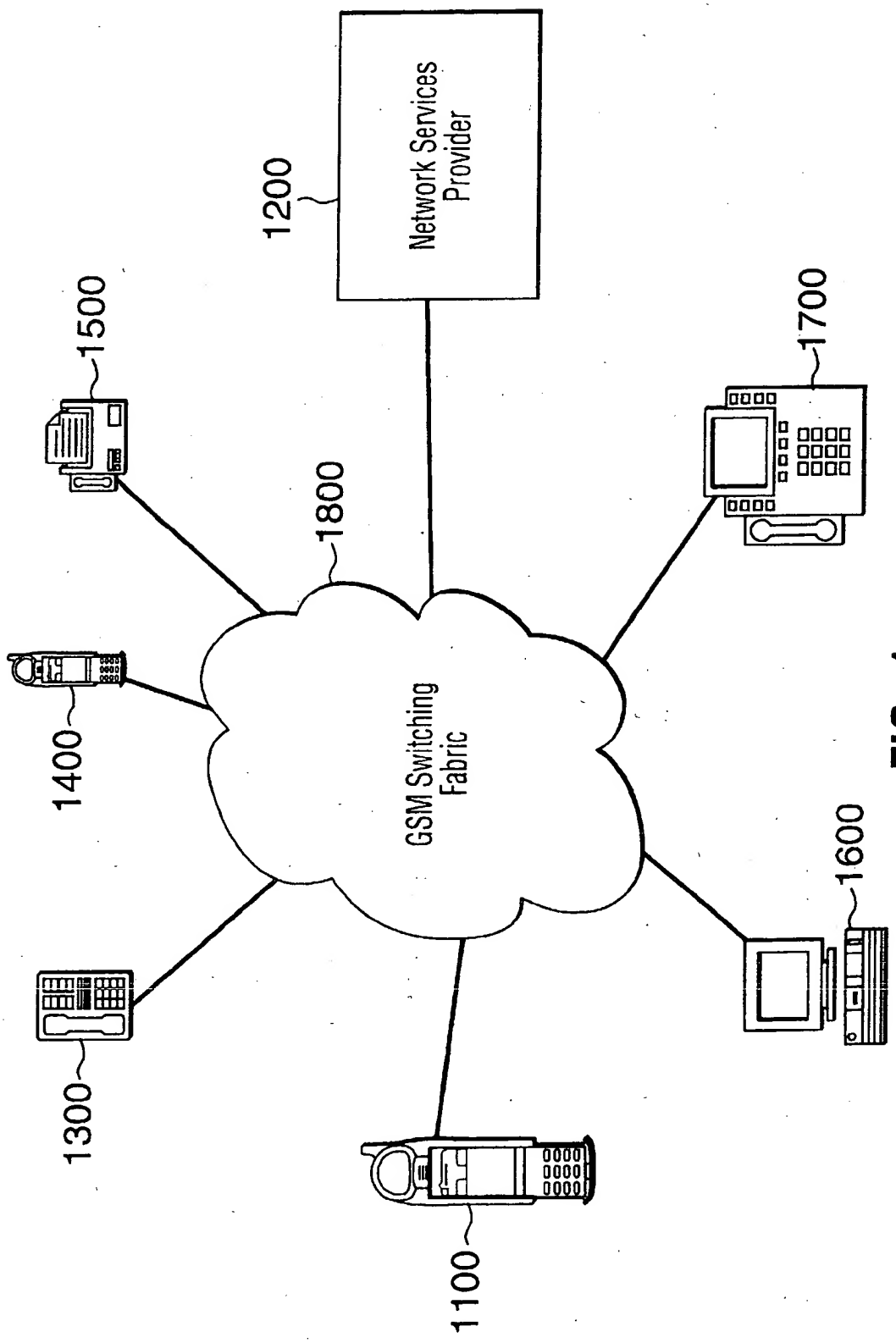

Network Services Provider
1200
1500
1700
1800
GSM Switching Fabric
1400
1600
1300
1100

FIG. 1

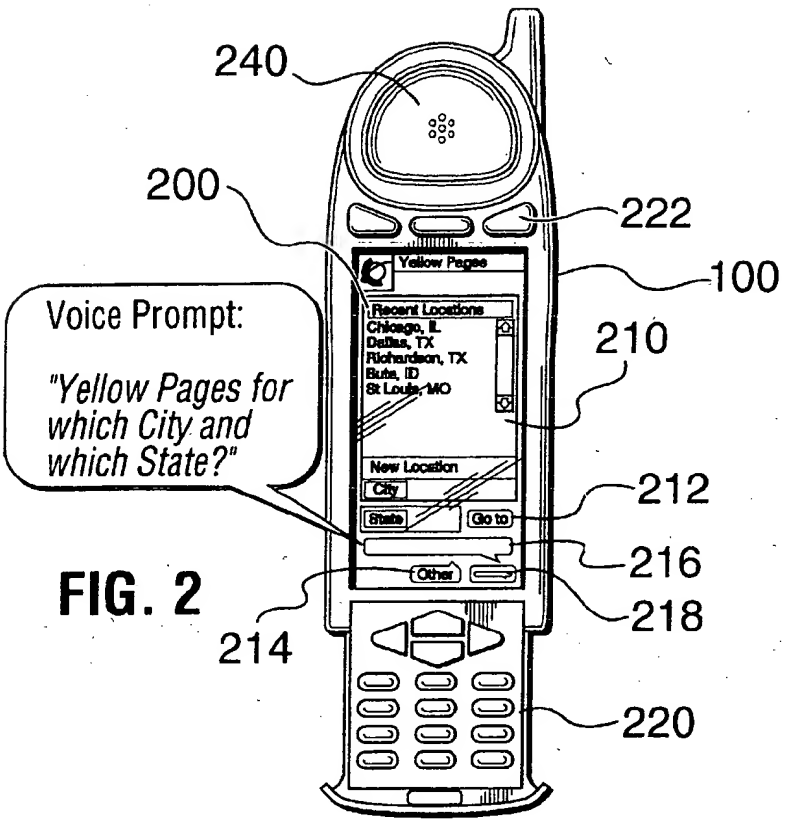

240
200
222
100
Voice Prompt:
"Yellow Pages for which City and which State?"
Yellow Pages
Recent Locations
Chicago, IL
Dallas, TX
Richardson, TX
Bute, ID
St Louis, MO
210
New Location
City
State
Go to
212
216
Other
218
214
220

FIG. 2

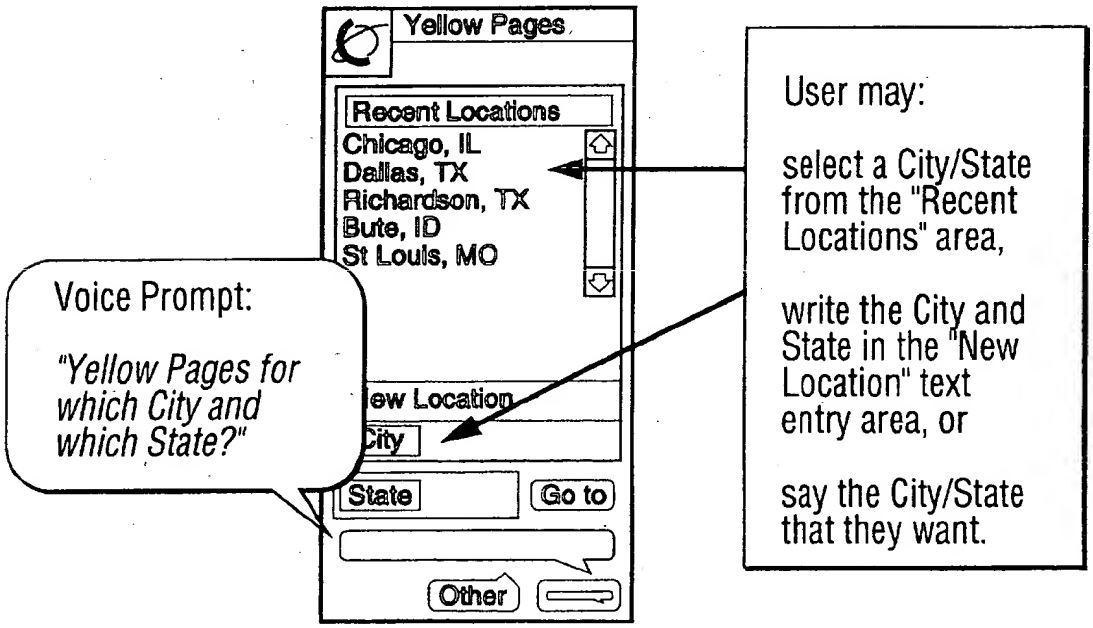

Yellow Pages
Recent Locations
Chicago, IL
Dallas, TX
Richardson, TX
Bute, ID
St Louis, MO
Voice Prompt:
"Yellow Pages for which City and which State?"
ew Location
City
State
Go to
Other
User may:
select a City/State from the "Recent Locations" area,
write the City and State in the "New Location" text entry area, or
say the City/State that they want.

FIG. 3

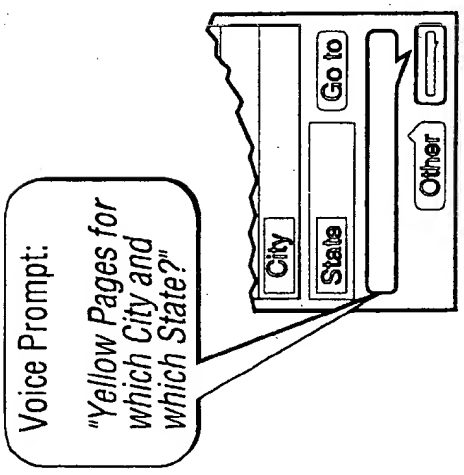


FIG. 4A

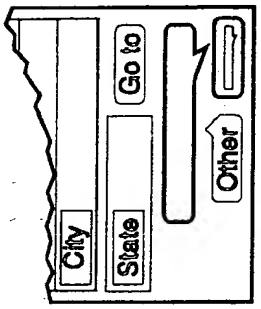


FIG. 4B

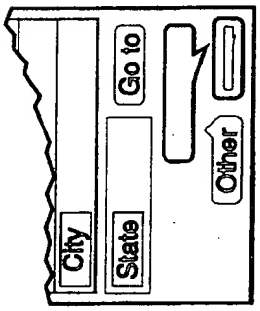


FIG. 4C

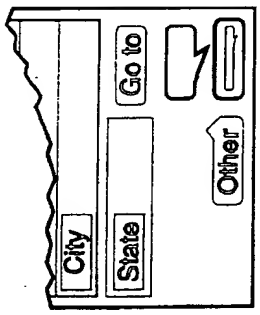


FIG. 4D

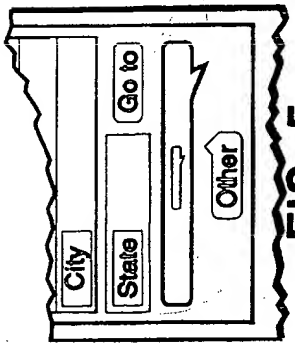


FIG. 5

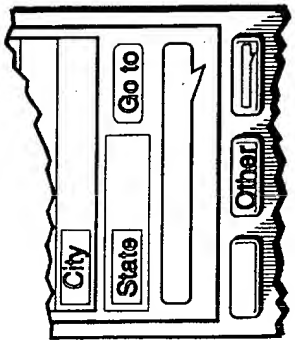


FIG. 6

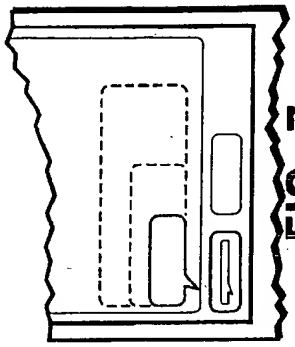


FIG. 7

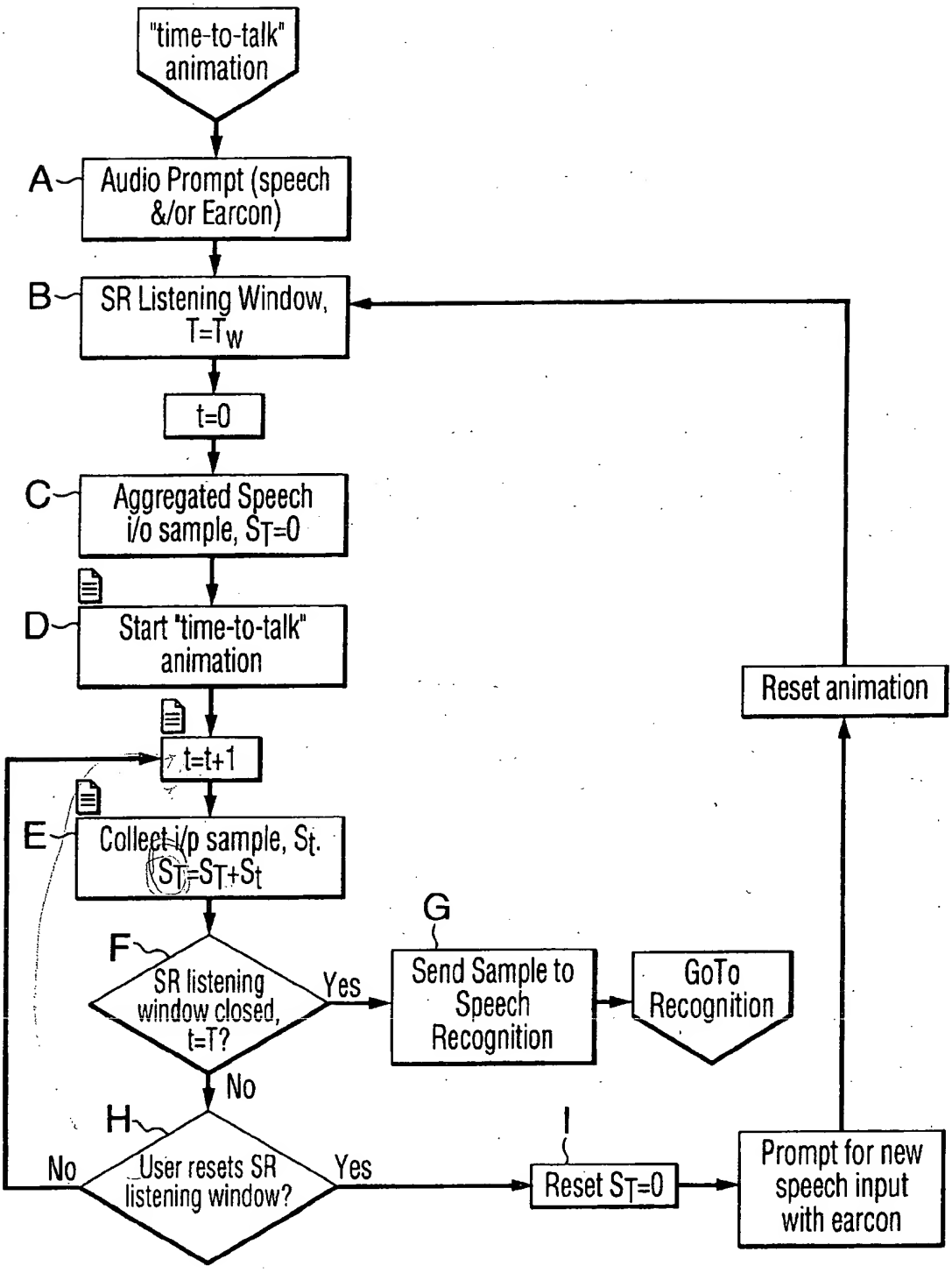

"time-to-talk" animation
A
Audio Prompt (speech &/or Earcon)
B
SR Listening Window, T=Tw
t=0
C
Aggregated Speech i/o sample, ST=0
D
Start "time-to-talk" animation
t=t+1
E
Collect i/p sample, St. ST=ST+St
F
SR listening window closed, t=T?
Yes
G
Send Sample to Speech Recognition
GoTo Recognition
No
H
User resets SR listening window?
No
Yes
I
Reset ST=0
Prompt for new speech input with earcon
Reset animation

FIG. 8

COMMUNICATION SYSTEM USER INTERFACE WITH ANIMATED REPRESENTATION OF TIME REMAINING FOR INPUT TO RECOGNIZER

RELATED APPLICATIONS

This application is related to U.S. patent application Ser. No. 09/062,969 entitled "Server for handling multimodal information" to H. Pasternak; and U.S. patent application Ser. No. 09/062,970 entitled 'Management of speech and audio prompts in multimodal interfaces' to M. French St. George, filed concurrently herewith.

FIELD OF THE INVENTION

This invention relates to management of interfaces for telecommunications systems and devices having a limited time window for capturing input, with particular application to speech recognizers having a limited time window for responding to a speech prompt.

BACKGROUND OF THE INVENTION

Telecommunications systems with a speech recognition capability have been in use for some time for performing basis tasks such as directory dialling. There are also network based speech recognition servers that deliver speech enabled directory dialling to any telephone. Typically, speech prompted interfaces have been used in telecommunications systems in contexts where there is no visual display or the user is unable to use visual displays, for example, a conventional telephone terminal.

Typically, the speech interface prompts the user when to speak by providing a speech prompt, i.e. a recognisable phrase or question prompting for user input, or by emitting a 'speak now' beep after which the speech recognizer is turned on, for a limited time window, typically a few seconds, during which the user may respond.

Users of telecommunications equipment employing speech recognition systems often report feeling rushed when prompted to respond immediately after a beep or other audible prompt.

Part of this rushed feeling may be attributed to a sense that the device will stop recognition before the user has completed their verbal request because the user receives no indication of the time window available to respond after the recognizer it turned on or when the recognition window is open. The user may find it difficult to know when the speech recognizer is on, may talk when the recognizer is off, or may become confused by no response.

Other difficulties may occur if the user does not remember what is acceptable input vocabulary or grammar to use. In addition to the sense of having to respond right now, current speech interface structures do not provide the user with an opportunity to rephrase a request, or change their mind before waiting for the system to respond. The user's utterance is accepted and interpreted, and the system advances to the next logical state, which may result in an error, for example if the user mis-speaks, coughs, or simply makes a mistake. Similarly, undue hesitation after a partial response may cause the system to move prematurely to the next logical state. If this state is not where the user wants to be, the user must navigate back to the previous state and restate the request.

Currently, the best recognizers in use have a 90 to 95 percent recognition performance under optimum conditions, and a noisy background environment, other speakers, user

accents, the user speaking to softly, may adversely affect recognition performance.

When conditions are not optimum, additional dialogue may assist. For example, the recognizer may give repeat instructions, or provide additional instructions. Nevertheless, using speech to provide additional information is slow. Consequently the user may perceive an excessively long wait for the system to reset and issue a new prompt. Typically, speech is perceived as fast for input, and slow for output.

Many users report becoming frustrated with using interactive voice response (IVR) systems offering many choices or a multi level menu system of choices. The user may forget long lists of instructions, or become confused or lost in a complex speech application.

User difficulties in interacting with these systems represent some reasons such speech interfaces have not yet gained as widespread acceptance as they might.

Older systems which also provide a graphical user interface, i.e. a screen display, with a speech interface, have been discrete non-integrated techniques That is the system may use either a touch input or a speech input, but not both simultaneously.

To overcome the inconvenience of switching between discrete applications offering different modes of interaction, systems are being developed to handle more than one type of interface, i.e. more than one mode of input and output, simultaneously. In the following description the term input/output modality refers to a sensory modality relating to a user's behaviour in interacting with the system, i.e. by using auditory, tactile and visual senses. Input/output modes refer to specific examples of use of these modalities. For example speech and audio input/output represent an auditory modality; use of a keypad, pen, and touch sensitive buttons represent a tactile input modality, and viewing a graphical display relies on the visual modality.

An example of a multimodal interface is 08/992,630 entitled "Multimodal User Interface", filed Dec. 19, 1997, to Smith and Beaton, which is incorporated herein by reference. This application discloses a multi-modal user interface and provides a telecommunications system and methods to facilitate multiple modes of interfacing with users for example, using voice, hard keys, touch sensitive soft key input, and pen input. This system provides, e.g. for voice or key input of data, and for graphical and speech data output. The user may choose to use the most convenient mode of interaction with the system and the system responds to input from all modes.

Thus, interfaces for communications devices and computer systems are becoming increasingly able to accept input and provide output by various modes.

For example, current speech recognition interfaces may be used in association with an visual display showing an icon that indicates current word recognition state. These icons change visually when the recognition state changes from listening to not-listening. For example, a "talk now" icon may be displayed in the corner of the screen. While these icons indicate to the user that the speech recognizer is on, the icons do not overcome the users perception of urgency to talk before the window closes. Also, as mentioned above if an error is made, or speech input is interrupted by extraneous background noise, the system waits until the 'talk now' or recognition window closes, and advances to the next logical state to recover from such an error, before issuing a new prompt and reopening the recognition window.

There also exist natural language speech interfaces that are always on, which preclude the need for beeps that inform the user of when to start talking. The user may speak at any time, and the recognizer will always be ready to listen. Currently this type of recognition is not yet widely distributed and used. These more advanced speech recognizers currently rely on a network based speech recognizer to provide the necessary processing power. Thus in the foreseeable future, this type of advanced speech recognition will co-exist with simpler forms of recognition that require a limited duration 'time to talk' window, or recognition window.

SUMMARY OF THE INVENTION

Thus, the present invention seeks to provide a system and method for management of an interface for a communications systems or devices including a limited time duration window for accepting input, and particularly for management of a speech interface for a speech recognizer with a limited duration 'time-to-talk' window, which avoids or reduces some of the above mentioned problems.

Therefore, according to a first aspect of the present invention there is provided a communications system comprising:

an interface operable for receiving input during limited time recognition window,

means for mapping available time of the recognition window to a spatial representation in animated form using one of a graphical modality, haptic modality or auditory modality.

For example the spatial representation comprises one of a two dimensional representation and three dimensional representation which diminishes in size as the available time diminishes, thereby providing a user with an indication of the available time to provide input.

Where the means for mapping comprises a graphical user interface, preferably the spatial representation comprises graphical information in the form of an animation, which diminishes in size as the available time diminishes. The rate of size change of the animation may be linear or non-linear, as required, to influence the users perception of the remaining time available.

Alternatively, the means for mapping comprises a haptic interface and the spatial representation comprises an animated pattern of haptic stimulation, e.g. produced by a wristband stimulator.

Alternatively, the means for mapping comprises an auditory interface the spatial representation comprises a time dependent auditory pattern.

These alternative modalities may be combined, for example to provide both graphical and haptic feedback to the user regarding available time in the recognition window.

Preferably, the system also provide a means for resetting the recognition window by the user, and reinitiating the animation. For example, the reset means may be a key or button associated with the animation. Alternatively, a touch sensitive region may be associated with the animation, e.g a touch sensitive button associated with the animation.

According to another aspect of the invention there is provided a communications device comprising:

an speech interface for accessing a speech recognizer operable for receiving speech input during limited time recognition window, and

means for mapping available time of the recognition window to a spatial representation in animated form using one of a graphical modality, haptic modality or auditory modality.

According to yet another aspect of the invention there is provided a communications system comprising:

a speech interface for accessing a speech recognizer operable for receiving speech input during limited time recognition window,

means for mapping available time of the recognition window to a spatial representation in animated form using one of a graphical modality, haptic modality or auditory modality.

Thus where the interface is a speech interface and the animation is an animation displayed on the graphical interface, the animation itself may provide the touch sensitive region which shrinks as the recognition window closes, and which is operable by the user to reset the 'time to talk' recognition window. An appropriate graphical animation for a speech interface is a speech balloon.

Preferably, the system comprises means for resetting the recognition window, which comprises a tactile input means associated with the animation displayed on the graphical user interface, e.g. a button.

For example when the animation is provided by a touch sensitive region of the graphical user interface, the means for resetting the recognition window comprises an associated touch sensitive region of the graphical user interface.

Where speech input has already been captured, the reset means also functions to delete speech input when the speech recognition window is reset.

According to another aspect of the present invention there is provided a method of providing user feedback and control for a communications device for accessing a recognizer for receiving input while the recognizer is operable during a limited time duration recognition, comprising the steps of:

after prompting the user for input;

turning on the recognizer for a limited time duration recognition window;

and while the recognizer is on, mapping available time to a spatial representation in animated form in one of a graphical modality, haptic modality and auditory modality.

Where the interface comprises a graphical user interface the method comprises providing a spatial representation in the form of an graphical animation indicative of the time remaining in the recognition window, and preferably comprises displaying an animation which diminishes in size as the time remaining in the recognition window decreases.

This method is particularly useful for communications device having a speech interface for accessing a speech recognizer, where the speech recognizer is operable for receiving input during a limited time duration recognition window.

Beneficially while the speech recognizer is on, the method comprises generating a timing signal representative of the time remaining in the recognition window,

controlling the animation using the timing signal and

at each time increment of the timing signal capturing speech input, and checking for user input to a reset means,

and when user input to the reset means is captured, resetting the recognition window, deleting captured speech input, and reinitiating the timing signal and the animation,

and when the timing signal corresponds to the end of the recognition window, capturing aggregated speech input and closing the recognition window.

Another aspect of the present invention provides software on a computer readable medium for carrying out these methods.

Conveniently, in use of a speech interface, the graphical information is an animation displayed on the graphical user interface which changes as the recognition window closes with time. For example the animation may be a speech balloon, or other appropriate icon, which diminishes in size as the recognition window closes. The animation provides the user with a graphical representation of the time remaining to talk. Movement on the screen also has the advantage of drawing the users attention to the closing recognition window.

Thus the user receives a visual indication of the time remaining in the recognition window of the speech recognizer, and is provided with the option to reset the recognition window, for example, to correct a speech input error. Thus the user receives feedback regarding the operation of the speech recognizer, and is provided with an opportunity to control opening of the recognition window, e.g. to correct errors.

Beneficially, by using a multitasking, multimodal user interface, and combining a touch sensitive display to the interface, an animated graphic can also serve as a reset button that the user may touch, to re-open a closing window.

This user behaviour resets the recognition window and allows the user to correct or alter a spoken command to a speech interface without having to wait for the system to achieve its next state before returning to correct an input. Reset means allows the user reset the time window at will. If not reset the animation stops when the time reaches the time duration Tw and the recognition window closes, after capturing aggregated speech input collected while the window is open.

Provision of an animated graphical output, with or without haptic and auditory feedback, indicative of time available, provides the user with additional timing information. Consequently user feedback and control of the interface or speech interface is improved.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention will now be described in greater detail with reference to the attached drawings wherein:

FIG. **1** shows a schematic block diagram of a communications network comprising a mobile telephone having a multitasking graphical user interface consistent with an embodiment of the present invention;

FIG. **2** shows schematically a diagram of mobile telephone of a first embodiment for operation in the network of FIG. **1**;

FIG. **3** shows part of the mobile telephone of FIG. **2** showing on an enlarged scale detail of the touch sensitive graphical display with pen input capability, during operation of a Yellow Pages directory application using a method according to an embodiment of the invention.

FIGS. **4A**, **4B**, **4C**, and **4D** show schematically part of the display shown in FIG. **3**, showing on an enlarged scale, to show the time to talk indicator and controller displayed by the graphical interface as the time window of the speech recognizer advances;

FIG. **5** shows schematically part of a graphical user interface representing an second embodiment of the indicator and controller;

FIG. **6** shows schematically part of a graphical user interface representing an third embodiment of the indicator and controller;

FIG. **7** shows schematically part of a graphical user interface representing an fourth embodiment of the indicator and controller;

FIG. **8** shows a flow chart setting out the steps of a method of providing interface management according to an embodiment of the present invention, i.e. user feedback and control of a speech interface using a multimodal interface of the mobile telephone shown in FIGS. **2** and **3**.

DETAILED DESCRIPTION OF THE INVENTION

A schematic block diagram of a communications network **10** is shown in FIG. **1** and represents a GSM switching services fabric **20** and a network services provider **40** associated with a plurality of communications terminals, for example a mobile telephone **100**, and other wired or wireless communications devices and terminals represented schematically by units **110**, **120**, **130**, **140** and **150**.

The wireless mobile phone **100** according to a first embodiment of the present invention is shown enlarged in FIG. **2**, and is provided with a multitasking graphical user interface, similar to that described in copending U.S. application Ser. No. entitled "Multimodal User Interface" filed Dec. 19, 1997, 08/992,630 to Smith and Beaton, which is incorporated herein by reference. This is a multi-modal user interface which provides a telecommunications system and methods to facilitate multiple modes of interfacing with users for example, using voice, hard keys, touch sensitive soft key input, and pen input. This system provides, e.g. for voice or key input of data, and for graphical and speech data output. Thus the user may choose to use the most convenient mode of interaction with the system and the system responds to input from all modes.

As shown in FIG. **2**, mobile telephone unit **100** comprises a body **200** which carries a display screen **210** for the graphical user interface, which may include touch sensitive buttons **212**; a conventional keypad **220** and other hard keys **222**; speaker **240** associated with the speech interface to providing speech prompts as shown schematically by **230** to illustrate the various modes of interaction which may be selected by a user. Thus the system provides for conventional key input, a graphical user interface which includes touch sensitive soft keys, and a speech interface for outputting speech and audio prompts, and a speech recognizer to accept and interpret speech input. Optionally, the graphical user interface may support pen input to allow written input on a touch sensitive area of the screen.

The device **100** according to the first embodiment of the present invention comprises means for mapping available time of the recognition window to a spatial representation in one of a graphical, haptic or auditory modality. This embodiment is an example using a graphical modality for providing user feedback and control of a 'time to talk' recognition window of the speech interface. In FIG. **2**, and in the enlarged schematic of the screen display shown in FIG. **3**, near the bottom of the display screen **210**, the user feedback means is an indicator which takes the form of an appropriate animated icon, or animation **216**, as will be described below, and is associated with a touch sensitive button **218**. The button **218** functions to control and reset the speech recognizer. The lower part of the screen **210** is shown enlarged in FIGS. **4A**, **4B**, **4C** and **4D** to show how the animation changes, i.e. shrinks, as the time to talk window of the speech recognizer is opened and closes.

Since speech and word recognition systems have limited recognition windows, when the speech recognizer is turned on and a recognition window is opened, a timing signal is sent to the device that starts the indicator, for example displaying the animation shown on a graphical display

represented in FIGS. 4A to 4D. The signal contains information regarding the timing of the animation, and as the recognition window advances and closes, the animation also changes and closes. In this example the animation is a speech balloon which diminishes in size as the time to talk window closes. The timing signal contains information regarding the timing of the animation. As the recognition window closes , the animation also closes.

The rate at which the animation diminishes in size is controlled by the timing signal. The rate of change of the animation may be linear or non-linear with respect to the time available. While the user does not receive a count down, i.e. it is not a numeric display of the time remaining, the animation provides a spatial representation of the time remaining, which allows the user to gauge how much time is left. The user receives sufficient visual feedback on timing information to provide and indication of the progress of the window, and warning when the recognition window will close. The rate of change of the animation may be used to influence the users perception of the available time and the rate at which the window is closing.

In the example shown in FIGS. **2**, **3** and 4A to 4D, the animation is a graphical speech balloon. Touching the button **218** associated with the animation, i.e. on an touch sensitive adjacent region of the screen display, sends a signal to the recognizer to delete any verbal signal associated with a previously opened recognition window in this state, and resets the timing signal, and the animation, indicating that the user may restate their request.

Thus, the user is provided with information regarding the time left in the recognition window, and an undo or reset feature is also provided. The reset feature allows users to correct speech utterances while the window is still open, or after closing the window, and before the system advances to an erroneous state.

Thus the speech interface is provided with animated graphical support, to indicate the time duration of an available recognition window.

Optionally, haptic support is provided in addition to graphical output, for example by providing a spatial representation of the time remaining through a haptic pattern generated by an array of stimulators carried on an associated wristband of the communications device.

While in this first embodiment, the balloon animation and balloon button are separate entities, advantageously, according to a second embodiment, the animation is provided on a touch sensitive area of the display, i.e. the animation itself is touch sensitive and functions as a soft button providing the reset function, as shown schematically in FIG. **5**.

Alternatively a soft key implementation of the indicator, i.e. the latter balloon button could alternatively is by a hard key button version, as shown schematically in FIG. **6**. The animation illustrated above shrinks in length as the available time diminishes, but the animation may alternatively take any convenient form, and inanother embodiment shown in FIG. **7**, the animation shrinks in 2 dimensions. These are simply given as examples, and any convenient form of animation may be used.

The user is provided with graphical and dynamic information regarding the time left in the recognition window.

The reset, or undo feature allows users to correct speech utterances before the system advances to an erroneous state.

Speech recognition may reside on the network, or at the terminal device, and the indicator and controller is applicable to speech and word recognition systems having limited recognition windows.

A method in accordance with the first embodiment of the present invention, with reference to the portable wireless terminal **100** shown in FIGS. **2** to **4**, will be described in more detail with reference to the flow chart shown in FIG. **8**.

In operation of the device, the use may pick up the device, and cause activation of the device either automatically upon picking up the device, or for example by pressing a button.

The speech recognizer is thereby switched on, and prompts the user for input (A). The prompt may be an audio, either by speech or an 'earcon', that is a recognizable sound such as a characteristic 'beep'. The speech recognizer (SR) window is turned on (B) for a specific recognition window $T=T_w$, and a timing signal is initiated at T=O, which is indicative of the time remaining, Tw, in the recognition window.

At time T=O, an aggregated speech input sample $S_I=O$ is captured (C) and a time-to-talk animation associated with the timing signal is displayed on the screen and the animation is started (D). The speech balloon is representing the speech recognition window, or 'time to talk' window, is displayed at its maximum size as shown in FIG. 4A.

At each specified time interval or time increment, t, an aggregated speech input sample is collected $S_I=S_O+S_r$. After each speech input is sampled, a check is made to determine if the speech recognizer window is open, i.e. if $T \leq Tw$.

When the time value is checked, if the speech recognizer recognition window has closed at T=Tw the aggregated speech sample is sent for speech recognition (G).

If the time window is still open T<Tw, the recognizer checks whether the user has issued a reset command, i.e. pressed the reset button, to reset the recognition window (H). If not, the timing signal advances, the collection of the next input speech sample is commenced (E). The animation continues to advance, diminishing in size continuously as the timing signal advances (see FIGS. 4B to 4D), to provide the user with a further visual indication of the diminishing available time to talk left in the recognition window.

When T=Tw, the aggregated speech sample is captured, the recognition is closed, i.e. the recognizer is turned off, and the animation ceases.

If at any time during the recognition window, the user resets the speech recognizer, any aggregated speech input that has been captured is erased, and a prompt for new speech input is issued, e.g. using a simple 'earcon'. The timing signal is reset, (I) and the animation is reset to the time zero display as shown in FIG. 4A.

The method of operation of the indicator and controller is generally applicable to all devices that access speech recognition engines, whether these are implemented locally, or on a network server.

The system and method described above was developed particularly to facilitate interaction with a speech recognizer. The system and method in general is not limited to speech interfaces, but may also be applied to other interfaces operating with a limited time window for receiving input. Another example may be a system for receiving and recognizing security information e.g. user id and password information, to provide access to an information system, wherein a time limit for access is provided, for example to deter unauthorized entry.

In another example (not shown) the haptic animation mentioned above may be provided instead of graphical information regarding the available time in the recognition window. In another example (not shown), the animation is

auditory, i.e. a pattern of sound representative of the available time may be generated by stereo speakers of the communications device.

While the systems and methods described above may be implemented by any convenient software and hardware configuration for local device or network server implementation of a speech interface, and the method is intended to be platform independent. For Internet compatible communications devices, the current application is a preferred implementation of the current application uses a Java web server, such as one designed for this application as described in the above referenced copending U.S. patent application Ser. No. 09/062,969, entitled "Server for handling multimodal information" to H. Pasternak filed concurrently herewith, and incorporated herein by reference. This server was developed for handling information in different modal forms associated with respective input/output modalities of a multimodal interface, and may be used to implement the systems and methods described herein.

Although specific embodiments of the invention have been described in detail, it will be apparent to one skilled in the art that variations and modifications to the embodiments may be made within the scope of the following claims.

What is claimed is:

1. A communications system comprising:

an interface operable for receiving input during limited time recognition window,

means for mapping available time of the recognition window to a spatial representation in animated form using one of a graphical modality, haptic modality or auditory modality.

2. A system according to claim 1 wherein the spatial representation comprises one of a two dimensional representation and three dimensional representation which diminishes in size as the available time diminishes.

3. A system according to claim 1 wherein the means for mapping comprises a graphical user interface and the spatial representation comprises graphical information in the form of an animation.

4. A system according to claim 3 wherein the animation diminishes in size as the available time diminishes.

5. A system according to claim 4 wherein the rate of change of the animation is linear.

6. A system according to claim 4 wherein the rate of change of animation is non-linear.

7. A system according to claim 1 wherein the means for mapping comprises a haptic interface and the spatial representation comprises an animated pattern of haptic stimulation.

8. A system according to claim 1 wherein the means for mapping comprises an auditory interface the spatial representation comprises an time dependent auditory pattern.

9. A system according to claim 1 comprising reset means for resetting the recognition window and reinitiating the animation.

10. A communications device comprising:

an speech interface for accessing a speech recognizer operable for receiving speech input during limited time recognition window, and

means for mapping available time of the recognition window to a spatial representation in animated form using one of a graphical modality, haptic modality or auditory modality.

11. A communications system comprising:

a speech interface for accessing a speech recognizer operable for receiving speech input during limited time recognition window,

means for mapping available time of the recognition window to a spatial representation in animated form using one of a graphical modality, haptic modality or auditory modality.

12. A system according to claim 11 wherein the means for mapping comprises a graphical user interface and the spatial representation comprises a graphical animation.

13. A communications system according to claim 12 comprising means for resetting the recognition window.

14. A communications system according to claim 13 wherein the means for resetting the recognition window comprises a tactile input means associated with the animation displayed on the graphical user interface.

15. A communications system according to claim 13 wherein the animation is provided by a touch sensitive region of the graphical user interface, and the means for resetting the recognition window comprises an associated touch sensitive region of the graphical user interface.

16. A system according to claim 11 wherein the spatial representation comprises one of a two dimensional representation and three dimensional representation which diminishes in size as the available time diminishes.

17. A system according to claim 11 wherein the means for mapping comprises a haptic interface and the spatial representation comprises an animated pattern of haptic stimulation.

18. A system according to claim 11 wherein the animation diminishes in size as the available time diminishes.

19. A system according to claim 18 wherein the rate of change in size of the animation is linear.

20. A system according to claim 18 wherein the rate of change in size of animation is non-linear.

21. A system according to claim 11 wherein the animation is indicative of the operation of the recognition window, and the animation changes during the operation of the recognition window to indicate the remaining time duration of the recognition window.

22. A method of providing user feedback and control for a communications device for accessing a recognizer for receiving input while the recognizer is operable during a limited time duration recognition, comprising the steps of:

after prompting the user for input;

turning on the recognizer for a limited time duration recognition window;

and while the speech recognizer is on, mapping available time to a spatial representation in animated form in one of a graphical modality, haptic modality and auditory modality.

23. A method according to claim 22 wherein the interface comprises a graphical user interface and the method comprises providing a spatial representation in the form of an graphical animation indicative of the time remaining in the recognition window.

24. A method according to claim 23 comprising displaying an animation which diminishes in size as the time remaining in the recognition window decreases.

25. A method of providing user feedback and control for a communications device having a speech interface for accessing a speech recognizer, the speech recognizer being operable for receiving input during a limited time duration recognition window, comprising the steps of:

after prompting the user for input;

turning on the recognizer for a limited time duration recognition window;

and while the speech recognizer is on, mapping available time to a spatial representation in animated form using

one of a graphical modality, haptic modality and auditory modality.

26. A method according to claim 25 wherein the interface comprises a graphical user interface and the method comprises providing a spatial representation in the form of an graphical animation indicative of the time remaining in the recognition window.

27. A method according to claim 26 comprising displaying an animation which diminishes in size as the time remaining in the recognition window decreases.

28. A method according to claim 27 comprising,

while the speech recognizer is on, generating a timing signal representative of the time remaining in the recognition window,

controlling the animation using the timing signal and

at each time increment of the timing signal capturing speech input, and checking for user input to a reset means,

and when user input to the reset means is captured, resetting the recognition window, deleting captured speech input, and reinitiating the timing signal and the animation,

and when the timing signal corresponds to the end of the recognition window, capturing aggregated speech input and closing the recognition window.

29. Software on a computer readable medium for carrying out a method of providing user feedback and control for a communications device having a speech interface for accessing a speech recognizer, the speech recognizer being operable for receiving input during a limited time duration recognition window, comprising the steps of:

after prompting the user for input;

turning on the recognizer for a limited time duration recognition window;

and while the speech recognizer is on, mapping available time to a spatial representation in animated form using one of a graphical modality, haptic modality and auditory modality.

* * * * *

APPENDIX E

(copies of cases cited)

signs was not legally erroneous, and because we find that the Board's finding that Valu's guide rails are *de jure* functional is supported by substantial evidence, the Board's refusal to register Valu's guide rail designs is *affirmed*, and Rexnord's cross-appeal is dismissed as moot.

AFFIRMED

COSTS

No costs.

In re Lee

U.S. Court of Appeals
Federal Circuit

No. 00-1158

Decided January 18, 2002

PATENTS

[1] Practice and procedure in Patent and Trademark Office — Board of Patent Appeals and Interferences — In general (§ 110.1101)

Patentability/Validity — Obviousness — Combining references (§ 115.0905)

Patentability/Validity — Obviousness — Evidence of (§ 115.0906)

Rejection of patent application for obviousness under 35 U.S.C. § 103 must be based on evidence comprehended by language of that section, and search for and analysis of prior art includes evidence relevant to finding of whether there is teaching, motivation, or suggestion to select and combine references relied on as evidence of obviousness; factual inquiry whether to combine references must be thorough and searching, based on objective evidence of record, and Board of Patent Appeals and Interferences must explain reasons why one of ordinary skill in art would have been motivated to select references and to combine them to render claimed invention obvious.

[2] Patentability/Validity — Obviousness — Combining references (§ 115.0905)

JUDICIAL PRACTICE AND PROCEDURE

Procedure — Judicial review — Standard of review — Patents (§ 410.4607.09)

Board of Patent Appeals and Interferences improperly relied upon "common knowledge and common sense" of person of ordinary skill in art to find invention of patent application obvious over combination of two prior art references, since factual question of motivation to select and combine references is material to patentability, and could not be resolved on subjective belief and unknown authority, since deferential review of agency decisions under Administrative Procedure Act reinforces obligation of board to develop evidentiary basis for its findings, since board's rejection of need for any specific hint or suggestion in particular reference to support combination constituted omission of relevant factor required by precedent, and thus was both legal error and arbitrary agency action, since board's findings must extend to all material facts and be documented on record, and since "common knowledge and common sense" are not specialized knowledge and expertise of agency contemplated by APA, and may not be substituted for evidence, although they may be applied to analysis of evidence.

PATENTS

[3] Practice and procedure in Patent and Trademark Office — Board of Patent Appeals and Interferences — In general (§ 110.1101)

Patentability/Validity — Obviousness — Evidence of (§ 115.0906)

JUDICIAL PRACTICE AND PROCEDURE

Procedure — Judicial review — Standard of review — Patents (§ 410.4607.09)

Patent examiners and Board of Patent Appeals and Interferences, in relying on what they assert to be general knowledge to negate patentability on ground of obviousness, must articulate that knowledge and place it on record, since examiners and board are pre-

sumed to act from viewpoint of person of ordinary skill in art in finding relevant facts, assessing significance of prior art, and making ultimate determination of obviousness issue; failure to do so is not consistent with either effective administrative procedure or effective judicial review, and board cannot rely on conclusory statements when dealing with particular combinations of prior art and specific claims, but must set forth rationale on which it relies.

[4] Procedure — Court of Appeals for the Federal Circuit (§ 410.03)

Procedure — Judicial review — Standard of review — Patents (§ 410.4607.09)

U.S. Court of Appeals for the Federal Circuit will not consider proposed alternative grounds for affirming decision of Board of Patent Appeals and Interferences rejecting patent application for obviousness, since alternative grounds were made at oral argument and constitute post hoc rationalization for agency action, consideration of which would deprive aggrieved party of fair opportunity to support its position.

Appeal from the U.S. Patent and Trademark Office, Board of Patent Appeals and Interferences.

Patent application of Sang-Su Lee, serial no. 07/631,210, directed to method of automatically displaying functions of video display device and demonstrating how to select and adjust functions to facilitate user response. Applicant appeals from decision upholding rejection of all claims for obviousness, and from reaffirmation of that decision on reconsideration. Reversed and remanded.

Richard H. Stern and Robert E. Bushnell, Washington, D.C., for Sang Su Lee.

Sidney O. Johnson Jr., associate solicitor, John M. Whealan, solicitor, and Raymond T. Chen, Maximilian R. Peterson, and Mark Nagumo, associate solicitors, Arlington, Va., for Director of U.S. Patent and Trademark Office.

Before Newman, Clevenger, and Dyk, circuit judges.

Newman, J.

Sang-Su Lee appeals the decision of the Board of Patent Appeals and Interferences of the United States Patent and Trademark Office, rejecting all of the claims of Lee's patent application Serial No. 07/631,210 entitled "Self-Diagnosis and Sequential-Display Method of Every Function." [1] We vacate the Board's decision for failure to meet the adjudicative standards for review under the Administrative Procedure Act, and remand for further proceedings.

The Prosecution Record

Mr. Lee's patent application is directed to a method of automatically displaying the functions of a video display device and demonstrating how to select and adjust the functions in order to facilitate response by the user. The display and demonstration are achieved using computer-managed electronics, including pulse-width modulation and auto-fine-tuning pulses, in accordance with procedures described in the specification. Claim 10 is representative:

> 10. A method for automatically displaying functions of a video display device, comprising:
>
> determining if a demonstration mode is selected;
>
> if said demonstration mode is selected, automatically entering a picture adjustment mode having a picture menu screen displaying a list of a plurality of picture functions; and
>
> automatically demonstrating selection and adjustment of individual ones of said plurality of picture functions.

The examiner rejected the claims on the ground of obviousness, citing the combination of two references: United States Patent No. 4,626,892 to Nortrup, and the Thunderchopper Helicopter Operations Handbook for a video game. The Nortrup reference describes a television set having a menu display by which the user can adjust various picture and audio functions; however, the Nortrup display does not include a demonstration of how to adjust the functions. The Thunderchopper Handbook describes the Thunderchopper game's video display as having a "demonstration mode" showing how to play the game; however, the Thunderchopper Handbook makes no mention of the adjustment of picture or audio functions. The examiner held that it

[1] *Ex parte Lee*, No. 1994-1989 (Bd. Pat. App. & Int. Aug. 30, 1994; on reconsid'n Sept. 29, 1999).

would have been obvious to a person of ordinary skill to combine the teachings of these references to produce the Lee system.

Lee appealed to the Board, arguing that the Thunderchopper Handbook simply explained how to play the Thunderchopper game, and that the prior art provided no teaching or motivation or suggestion to combine this reference with Nortrup, or that such combination would produce the Lee invention. The Board held that it was not necessary to present a source of a teaching, suggestion, or motivation to combine these references or their teachings. The Board stated:

> The conclusion of obviousness may be made from common knowledge and common sense of a person of ordinary skill in the art without any specific hint or suggestion in a particular reference.

Board op. at 7. The Board did not explain the "common knowledge and common sense" on which it relied for its conclusion that "the combined teachings of Nortrup and Thunderchopper would have suggested the claimed invention to those of ordinary skill in the art."

Lee filed a request for reconsideration, to which the Board responded after five years. The Board reaffirmed its decision, stating that the Thunderchopper Handbook was "analogous art" because it was "from the same field of endeavor" as the Lee invention, and that the field of video games was "reasonably pertinent" to the problem of adjusting display functions because the Thunderchopper Handbook showed video demonstrations of the "features" of the game. On the matter of motivation to combine the Nortrup and Thunderchopper references, the Board stated that "we maintain the position that we stated in our prior decision" and that the Examiner's Answer provided "a well reasoned discussion of why there is sufficient motivation to combine the references." The Board did not state the examiner's reasoning, and review of the Examiner's Answer reveals that the examiner merely stated that both the Nortrup function menu and the Thunderchopper demonstration mode are program features and that the Thunderchopper mode "is user-friendly" and it functions as a tutorial, and that it would have been obvious to combine them.

Lee had pressed the examiner during prosecution for some teaching, suggestion, or motivation in the prior art to select and combine the references that were relied on to show obviousness. The Examiner's Answer before the Board, plus a Supplemental Answer, stated that the combination of Thunderchopper with Nortrup "would have been obvious to one of ordinary skill in the art since the demonstration mode is just a programmable feature which can be used in many different device[s] for providing automatic introduction by adding the proper programming software," and that "another motivation would be that the automatic demonstration mode is user friendly and it functions as a tutorial." The Board adopted the examiner's answer, stating "the examiner has provided a well reasoned discussion of these references and how the combination of these references meets the claim limitations." However, perhaps recognizing that the examiner had provided insufficient justification to support combining the Nortrup and Thunderchopper references, the Board held, as stated *supra*, that a "specific hint or suggestion" of motivation to combine was not required.

This appeal followed.

Judicial Review

Tribunals of the PTO are governed by the Administrative Procedure Act, and their rulings receive the same judicial deference as do tribunals of other administrative agencies. *Dickinson v. Zurko*, 527 U.S. 150, 50 USPQ2d 1930 (1999). Thus on appeal we review a PTO Board's findings and conclusions in accordance with the following criteria:

> 5 U.S.C. § 706(2) The reviewing court shall—
>
> (2) hold unlawful and set aside agency actions, findings, and conclusions found to be—
>
> (A) arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law;
>
> * * * *
>
> (E) unsupported by substantial evidence in a case subject to sections 556 and 557 of this title or otherwise reviewed on the record of an agency hearing provided by statute;

For judicial review to be meaningfully achieved within these strictures, the agency tribunal must present a full and reasoned explanation of its decision. The agency tribunal

must set forth its findings and the grounds thereof, as supported by the agency record, and explain its application of the law to the found facts. The Court has often explained:

> The Administrative Procedure Act, which governs the proceedings of administrative agencies and related judicial review, establishes a scheme of "reasoned decisionmaking." Not only must an agency's decreed result be within the scope of its lawful authority, but the process by which it reaches that result must be logical and rational.

Allentown Mack Sales and Service, Inc. v. National Labor Relations Bd., 522 U.S. 359, 374 (1998) (citation omitted). This standard requires that the agency not only have reached a sound decision, but have articulated the reasons for that decision. The reviewing court is thus enabled to perform meaningful review within the strictures of the APA, for the court will have a basis on which to determine "whether the decision was based on the relevant factors and whether there has been a clear error of judgment." *Citizens to Preserve Overton Park v. Volpe*, 401 U.S. 402, 416 (1971). Judicial review of a Board decision denying an application for patent is thus founded on the obligation of the agency to make the necessary findings and to provide an administrative record showing the evidence on which the findings are based, accompanied by the agency's reasoning in reaching its conclusions. *See In re Zurko*, 258 F.3d 1379, 1386, 59 USPQ2d 1693, 1697 (Fed. Cir. 2001) (review is on the administrative record); *In re Gartside*, 203 F.3d 1305, 1314, 53 USPQ2d 1769, 1774 (Fed. Cir. 2000) (Board decision "must be justified within the four corners of the record").

[1] As applied to the determination of patentability *vel non* when the issue is obviousness, "it is fundamental that rejections under 35 U.S.C. § 103 must be based on evidence comprehended by the language of that section." *In re Grasselli*, 713 F.2d 731, 739, 218 USPQ 769, 775 (Fed. Cir. 1983). The essential factual evidence on the issue of obviousness is set forth in *Graham v. John Deere Co.*, 383 U.S. 1, 17-18, 148 USPQ 459, 467 (1966) and extensive ensuing precedent. The patent examination process centers on prior art and the analysis thereof. When patentability turns on the question of obviousness, the search for and analysis of the prior art includes evidence relevant to the finding of whether there is a teaching, motivation, or suggestion to select and combine the references relied on as evidence of obviousness. *See, e.g., McGinley v. Franklin Sports, Inc.*, 262 F.3d 1339, 1351-52, 60 USPQ2d 1001, 1008 (Fed. Cir. 2001) ("the central question is whether there is reason to combine [the] references," a question of fact drawing on the *Graham* factors).

"The factual inquiry whether to combine references must be thorough and searching." *Id.* It must be based on objective evidence of record. This precedent has been reinforced in myriad decisions, and cannot be dispensed with. *See, e.g., Brown & Williamson Tobacco Corp. v. Philip Morris Inc.*, 229 F.3d 1120, 1124-25, 56 USPQ2d 1456, 1459 (Fed. Cir. 2000) ("a showing of a suggestion, teaching, or motivation to combine the prior art references is an 'essential component of an obviousness holding' ") (quoting *C.R. Bard, Inc., v. M3 Systems, Inc.*, 157 F.3d 1340, 1352, 48 USPQ2d 1225, 1232 (Fed. Cir. 1998)); *In re Dembiczak*, 175 F.3d 994, 999, 50 USPQ2d 1614, 1617 (Fed. Cir. 1999) ("Our case law makes clear that the best defense against the subtle but powerful attraction of a hindsight-based obviousness analysis is rigorous application of the requirement for a showing of the teaching or motivation to combine prior art references."); *In re Dance*, 160 F.3d 1339, 1343, 48 USPQ2d 1635, 1637 (Fed. Cir. 1998) (there must be some motivation, suggestion, or teaching of the desirability of making the specific combination that was made by the applicant); *In re Fine*, 837 F.2d 1071, 1075, 5 USPQ2d 1596, 1600 (Fed. Cir. 1988) (" 'teachings of references can be combined *only* if there is some suggestion or incentive to do so.' ") (emphasis in original) (quoting *ACS Hosp. Sys., Inc. v. Montefiore Hosp.*, 732 F.2d 1572, 1577, 221 USPQ 929, 933 (Fed. Cir. 1984)).

The need for specificity pervades this authority. *See, e.g., In re Kotzab*, 217 F.3d 1365, 1371, 55 USPQ2d 1313, 1317 (Fed. Cir. 2000) ("particular findings must be made as to the reason the skilled artisan, with no knowledge of the claimed invention, would have selected these components for combination in the manner claimed"); *In re Rouffet*, 149 F.3d 1350, 1359, 47 USPQ2d 1453, 1459 (Fed. Cir. 1998) ("even when the level of skill in the art is high, the Board must identify specifically the principle, known to one of ordinary skill, that suggests the claimed combina-

tion. In other words, the Board must explain the reasons one of ordinary skill in the art would have been motivated to select the references and to combine them to render the claimed invention obvious."); *In re Fritch*, 972 F.2d 1260, 1265, 23 USPQ2d 1780, 1783 (Fed. Cir. 1992) (the examiner can satisfy the burden of showing obviousness of the combination "only by showing some objective teaching in the prior art or that knowledge generally available to one of ordinary skill in the art would lead that individual to combine the relevant teachings of the references").

[2] With respect to Lee's application, neither the examiner nor the Board adequately supported the selection and combination of the Nortrup and Thunderchopper references to render obvious that which Lee described. The examiner's conclusory statements that "the demonstration mode is just a programmable feature which can be used in many different device[s] for providing automatic introduction by adding the proper programming software" and that "another motivation would be that the automatic demonstration mode is user friendly and it functions as a tutorial" do not adequately address the issue of motivation to combine. This factual question of motivation is material to patentability, and could not be resolved on subjective belief and unknown authority. It is improper, in determining whether a person of ordinary skill would have been led to this combination of references, simply to "[use] that which the inventor taught against its teacher." *W.L. Gore v. Garlock, Inc.*, 721 F.2d 1540, 1553, 220 USPQ 303, 312-13 (Fed. Cir. 1983). Thus the Board must not only assure that the requisite findings are made, based on evidence of record, but must also explain the reasoning by which the findings are deemed to support the agency's conclusion.

Deferential judicial review under the Administrative Procedure Act does not relieve the agency of its obligation to develop an evidentiary basis for its findings. To the contrary, the Administrative Procedure Act reinforces this obligation. *See, e.g., Motor Vehicle Manufacturers Ass'n v. State Farm Mutual Automobile Ins. Co.*, 463 U.S. 29, 43 (1983) ("the agency must examine the relevant data and articulate a satisfactory explanation for its action including a 'rational connection between the facts found and the choice made.' ") (quoting *Burlington Truck Lines v. United States*, 371 U.S. 156, 168 (1962)); *Securities & Exchange Comm'n v. Chenery Corp.*, 318 U.S. 80, 94 (1943) ("The orderly function of the process of review requires that the grounds upon which the administrative agency acted are clearly disclosed and adequately sustained.").

In its decision on Lee's patent application, the Board rejected the need for "any specific hint or suggestion in a particular reference" to support the combination of the Nortrup and Thunderchopper references. Omission of a relevant factor required by precedent is both legal error and arbitrary agency action. *See Motor Vehicle Manufacturers*, 463 U.S. at 43 ("an agency rule would be arbitrary and capricious if the agency . . . entirely failed to consider an important aspect of the problem"); *Mullins v. Department of Energy*, 50 F.3d 990, 992 (Fed. Cir. 1995) ("It is well established that agencies have a duty to provide reviewing courts with a sufficient explanation for their decisions so that those decisions may be judged against the relevant statutory standards, and that failure to provide such an explanation is grounds for striking down the action."). As discussed in *National Labor Relations Bd. v. Ashkenazy Property Mgt. Corp.*, 817 F.2d 74, 75 (9th Cir. 1987), an agency is "not free to refuse to follow circuit precedent."

The foundation of the principle of judicial deference to the rulings of agency tribunals is that the tribunal has specialized knowledge and expertise, such that when reasoned findings are made, a reviewing court may confidently defer to the agency's application of its knowledge in its area of expertise. Reasoned findings are critical to the performance of agency functions and judicial reliance on agency competence. *See Baltimore and Ohio R. R. Co. v. Aberdeen & Rockfish R. R. Co.*, 393 U.S. 87, 91-92 (1968) (absent reasoned findings based on substantial evidence effective review would become lost "in the haze of so-called expertise"). The "common knowledge and common sense" on which the Board relied in rejecting Lee's application are not the specialized knowledge and expertise contemplated by the Administrative Procedure Act. Conclusory statements such as those here provided do not fulfill the agency's obligation. This court explained in *Zurko*, 258 F.3d at 1385, 59 USPQ2d at 1697, that "deficiencies of the cited references cannot be remedied by

the Board's general conclusions about what is 'basic knowledge' or 'common sense.' " The Board's findings must extend to all material facts and must be documented on the record, lest the "haze of so-called expertise" acquire insulation from accountability. "Common knowledge and common sense," even if assumed to derive from the agency's expertise, do not substitute for authority when the law requires authority. *See Allentown Mack*, 522 U.S. at 376 ("Because reasoned decisionmaking demands it, and because the systemic consequences of any other approach are unacceptable, the Board must be required to apply in fact the clearly understood legal standards that it enunciates in principle . . . .")

The case on which the Board relies for its departure from precedent, *In re Bozek*, 416 F.2d 1385, 163 USPQ 545 (CCPA 1969), indeed mentions "common knowledge and common sense," the CCPA stating that the phrase was used by the Solicitor to support the Board's conclusion of obviousness based on evidence in the prior art. *Bozek* did not hold that common knowledge and common sense are a substitute for evidence, but only that they may be applied to analysis of the evidence. *Bozek* did not hold that objective analysis, proper authority, and reasoned findings can be omitted from Board decisions. Nor does *Bozek*, after thirty-two years of isolation, outweigh the dozens of rulings of the Federal Circuit and the Court of Customs and Patent Appeals that determination of patentability must be based on evidence. This court has remarked, in *Smiths Industries Medical Systems, Inc. v. Vital Signs, Inc.*, 183 F.3d 1347, 1356, 51 USPQ2d 1415, 1421 (Fed. Cir. 1999), that *Bozek*'s reference to common knowledge "does not in and of itself make it so" absent evidence of such knowledge.

[3] The determination of patentability on the ground of unobviousness is ultimately one of judgment. In furtherance of the judgmental process, the patent examination procedure serves both to find, and to place on the official record, that which has been considered with respect to patentability. The patent examiner and the Board are deemed to have experience in the field of the invention; however, this experience, insofar as applied to the determination of patentability, must be applied from the viewpoint of "the person having ordinary skill in the art to which said subject matter pertains," the words of section 103. In finding the relevant facts, in assessing the significance of the prior art, and in making the ultimate determination of the issue of obviousness, the examiner and the Board are presumed to act from this viewpoint. Thus when they rely on what they assert to be general knowledge to negate patentability, that knowledge must be articulated and placed on the record. The failure to do so is not consistent with either effective administrative procedure or effective judicial review. The board cannot rely on conclusory statements when dealing with particular combinations of prior art and specific claims, but must set forth the rationale on which it relies.

Alternative Grounds

[4] At oral argument the PTO Solicitor proposed alternative grounds on which this court might affirm the Board's decision. However, as stated in *Burlington Truck Lines, Inc. v. United States*, 371 U.S. 156, 168 (1962), "courts may not accept appellate counsel's *post hoc* rationalization for agency action." Consideration by the appellate tribunal of new agency justifications deprives the aggrieved party of a fair opportunity to support its position; thus review of an administrative decision must be made on the grounds relied on by the agency. "If those grounds are inadequate or improper, the court is powerless to affirm the administrative action by substituting what it considers to be a more adequate or proper basis." *Securities & Exchange Comm'n v. Chenery Corp.*, 332 U.S. 194, 196 (1947). As reiterated in *Federal Election Comm'n v. Akins*, 524 U.S. 11, 25 (1998), "If a reviewing court agrees that the agency misinterpreted the law, it will set aside the agency's action and remand the case — even though the agency (like a new jury after a mistrial) might later, in the exercise of its lawful discretion, reach the same result for a different reason." Thus we decline to consider alternative grounds that might support the Board's decision.

Further Proceedings

Sound administrative procedure requires that the agency apply the law in accordance with statute and precedent. The agency tribunal must make findings of relevant facts, and present its reasoning in sufficient detail that the court may conduct meaningful review of the agency action. In *Radio-Television News Directors Ass'n v. FCC*, 184 F.3d 872 (D.C.

Cir. 1999) the court discussed the "fine line between agency reasoning that is 'so crippled as to be unlawful' and action that is potentially lawful but insufficiently or inappropriately explained," quoting from *Checkosky v. Securities & Exch. Comm'n*, 23 F.3d 452, 464 (D.C. Cir. 1994); the court explained that "[i]n the former circumstance, the court's practice is to vacate the agency's order, while in the latter the court frequently remands for further explanation (including discussion of the relevant factors and precedents) while withholding judgment on the lawfulness of the agency's proposed action." *Id.* at 888. In this case the Board's analysis of the Lee invention does not comport with either the legal requirements for determination of obviousness or with the requirements of the Administrative Procedure Act that the agency tribunal set forth the findings and explanations needed for "reasoned decisionmaking." Remand for these purposes is required. *See Overton Park*, 401 U.S. at 420-221 (remanding for further proceedings appropriate to the administrative process).

VACATED AND REMANDED

Barbour v. Head

U.S. District Court
Southern District of Texas

No. G-01-491

Decided December 21, 2001

COPYRIGHTS

[1] Non-copyrightable matter — Ideas and systems (§ 211.05)

Defendants are not entitled to summary judgment that plaintiffs' cooking recipes are uncopyrightable, even though 17 U.S.C. § 102(b) denies copyright protection to mere procedures or processes, since neither courts nor Register of Copyrights have declared that recipes are per se uncopyrightable, since defendants have not shown that plaintiffs' cookbook is copyrighted as factual compilation or collective work rather than literary work, and since even if book is not literary work, genuine issue of material fact exists as to whether plaintiffs' recipes, which contain more than mechanical listings of ingredients and cooking instructions, represent mere unprotected facts or protectable expression.

JUDICIAL PRACTICE AND PROCEDURE

[2] Procedure — Limitations period; timeliness (§ 410.05)

Plaintiffs' claim for copyright infringement is not barred by three-year statute of limitations specified by 17 U.S.C. § 507(b), even though infringement claim was brought more than three years after infringing work was first published, since discovery rule and other equitable tolling doctrines apply to copyright claims, since plaintiffs' cause of action arguably did not accrue until they discovered defendants' book, less than one year before suit was brought, and since even if claim accrued on date of first publication, limitations period bars only remedy, not substantive right.

Action by Judy Barbour and Cookbook Resources LLC against James Head and Penfield Press Inc. for copyright infringement, and for unfair competition through misappropriation and conversion. On defendants' motion for summary judgment. Denied as to copyright claims; granted as to state law claims.

G.P. Hardy III, Houston, Texas, for plaintiffs.

Karen Bryant Tripp, Houston, for defendants.

Kent, J.

ORDER GRANTING IN PART DEFENDANT PENFIELD PRESS' MOTION TO DISMISS

This case involves a rustled cowboy cookbook. On August 13, 2001, Plaintiffs Judy Barbour ("Barbour") and Cookbook Resources, L.L.C. ("Cookbook Resources") filed causes of action for copyright infringement, unfair competition through misappropriation, and conversion, with which they're fixin' to brand Defendants James Head ("Head") and Penfield Press, Inc. ("Penfield Press"). On October 25, 2001, to bust out of the corral, Defendant Penfield Press filed a Motion to Dismiss pursuant to Fed.R.Civ.P. 12(b)(6). For the reasons articulated below, Defendant's Motion to Dismiss shall be

mately 80% of the total amount sought for that period.

[6] A total award of $130,643.75 for attorneys' fees is appropriate in this case in light of the *Kimbrell's* factors that apply. Preparation for the trial with respect to this particular defendant involved several pretrial motions, several raising complex factual questions, and the fast pace of this docket required an intense dedication of attorney time to the litigation. (Factors 1, 2 and 7). Moreover, the Court notes that plaintiffs' counsel took the case on a contingent fee basis, which means that given the amount of the judgment, they will be made whole for all work expended in this case. We also point out that the actual trial was relatively short, partly as a result of the defendant's having so little evidence to present. Finally, in light of the damages awarded, $130.643.75 is not an unreasonably high attorneys' fee. (Factor 8 and 12).

With respect to costs, the Court finds it appropriate, for the same reasons discussed above, to award the plaintiffs the total amounts reflected only on the law firm's invoices dated March 26, 1998 ($37,612.84) and April 9, 1998 ($17,309.51), for a total award of $54,922.35. This amount includes all costs incurred from February 23, 1998, to April 9, 1998.

IV. *Conclusion*

For the reasons stated in open court and in this Memorandum Opinion, BRFG's Motion for Declaration of Mistrial and Motion for a JNOV or in the Alternative a New Trial will be denied. In addition, plaintiffs' Motion of Trebled Damages and Attorneys' Fees will be denied as to the trebled damages and granted as to the attorneys' fees and costs, and the Court will award plaintiffs $130,643.75 in attorneys' fees and $54,922.25 in costs. An appropriate order will issue.

The Clerk is directed to forward copies of this Memorandum Opinion to counsel of record.

ORDER

For the reasons stated in an accompanying Memorandum Opinion, SRFG's Motion for Declaration of Mistrial and Motion for a JNOV or in the Alternative a New Trial are DENIED, and plaintiffs' Motion for Trebled Damages and Attorneys' Fees is GRANTED IN PART as to attorneys' fees and costs and DENIED as to trebled damages, and it is hereby.

ORDERED that plaintiffs be and are awarded $130,643.75 in attorneys' fees and $54,922.35 in costs, for a total award of $185,566.10.

The Clerk is directed to forward copies of this Order to counsel of record.

U.S. Court of Appeals
Federal Circuit

In re Rouffet

No. 97-1492

Decided July 15, 1998

PATENTS

1. Patentability/Validity — Obviousness — Combining references (§115.0905)

Claimed low orbit satellite communications system for mobile terminals, which addresses problem of minimizing "handover" of receiver from beam footprint of one transmitting satellite to that of another through use of multiple fan-shaped beams, is not prima facie obvious over combination of three prior art references, since critical reference that teaches use of fan-shaped beam to transmit from ground station to orbiting satellites does not specifically address handover minimization, and to extent it addresses handover problem at all, does so with orbit selection rather than beam shape, and since there is no reason one of ordinary skill in art, seeking to minimize handovers due to satellite motion, would have been motivated to combine this reference with remaining references in manner that would render claimed invention obvious.

2. Patentability/Validity — Obviousness — Person of ordinary skill in art (§115.0902)

Patentability/Validity — Obviousness — Combining references (§115.0905)

Three possible sources for motivation to combine prior art references in manner that would render claimed invention obvious are nature of problem to be solved, teachings of prior art, and knowledge of persons of ordinary skill in art; high level of skill in field of art cannot be relied upon to provide necessary motivation absent explanation of what specific understanding or technical principle, within knowledge one of ordinary skill in art, would have suggested combination, since, if such rote invocation could suffice to supply motivation to combine, more sophisticated scientific fields would rarely, if ever, experience patentable technical advance.

3. Patentability/Validity — Obviousness — Person of ordinary skill in art (§115.0902)

Patentability/Validity — Obviousness — Combining references (§115.0905)

Claimed low orbit satellite communications system for mobile terminals is not prima facie obvious over combination of two prior art references, even though person possessing high level of skill characteristic of this field would know to account for differences between claimed invention and prior art combination, since high level of skill in art, without more, cannot supply required motivation to combine references, and does not overcome absence of any actual suggestion to combine; obviousness rejection will not be upheld, even where skill in art is high, absent specific identification of principle, known to one of ordinary skill, that suggests claimed combination.

Appeal from the U.S. Patent and Trademark Office, Board of Patent Appeals and Interferences.

Patent application of Denis Rouffet, Yannick Tanguy, and Frédéric Berthault, serial no. 07/888,791, filed May 27, 1992. From decision upholding examiner's final rejection of application as obvious under 35 USC 103(a), applicants appeal. Reversed.

Richard C. Turner and Grant K. Rowan, of Sughrue, Mion, Zinn, Macpeak & Seas, Washington, D.C., for appellants.

David J. Ball Jr., associate solicitor, Nancy J. Linck, solicitor, Albin F. Drost, deputy solicitor, Craig R. Kaufman, associate solicitor, and Scott A. Chambers, associate solicitor, U.S. Patent and Trademark Office, Arlington, Va., for appellee.

Before Plager, circuit, judge, Archer, senior circuit judge, and Rader, circuit judge.

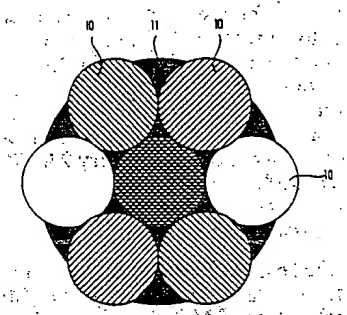
Rader, J.

Denis Rouffet, Yannick Tanguy, and Frédéric Berthault (collectively, Rouffet) submitted application 07/888,791 (the application) on May 27, 1992. The Board of Patent Appeals and Interferences (the Board) affirmed final rejection of the application as obvious under 35 U.S.C. § 103(a). *See Ex parte Rouffet*, No. 96-1553 (Bd. Pat. App. & Int. Apr. 16, 1997). Because the Board reversibly erred in identifying a motivation to combine the references, this court reverses.

I.

Satellites in a geosynchronous or geostationary orbit remain over the same point on the Earth's surface. Their constant position above the Earth's surface facilitates communications. These satellites project a number of beams to the Earth. Each beam transmits to its area of coverage, or footprint, on the Earth's surface. In order to provide complete coverage, adjacent footprints overlap slightly and therefore must use different frequencies to avoid interference. However, two or more non-overlapping footprints can use the same set of frequencies in order to use efficiently the limited radio spectrum. Figure 1 from the application shows the coverage of a portion of the Earth's surface provided by multiple cone shaped beams:

FIG. 1

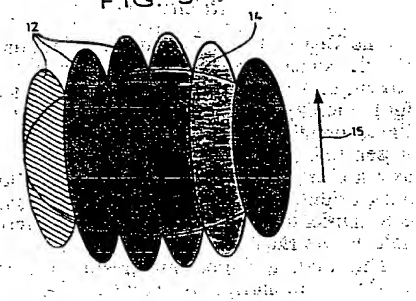


Frequency reuse techniques, however, have a limited ability to compensate for congestion in geostationary orbits. To alleviate the orbit congestion problem, new telecommunications systems use a network of satellites in low Earth orbit. When viewed from a fixed point on the Earth's surface, such satellites do not remain stationary but move overhead. A satellite's motion as it transmits a plurality of cone-shaped beams creates a new problem. The satellite's movement causes a receiver on the Earth's surface to move from the footprint of one beam into a second beam transmitted by the same satellite. Eventually, the satellite's motion causes the receiver to move from the footprint of a beam transmitted by one satellite into the footprint of a beam transmitted by a second satellite. Each switch from one footprint to another creates a "handover" event analogous to that which occurs when a traditional cellular phone travels from one cell to another. Handovers are undesirable because

they can cause interruptions in signal transmission and reception.

Rouffet's application discloses technology to reduce the number of handovers between beams transmitted by the same satellite. In particular, Rouffet eliminates handovers caused solely by the satellite's motion. To accomplish this goal, Rouffet changes the shape of the beam transmitted by the satellite's antenna. Rouffet's satellites transmit fan-shaped beams. A fan beam has an elliptical footprint. Rouffet aligns the long axis of his beams parallel to the direction of the satellite's motion across the Earth's surface. By elongating the beam's footprint in the direction of satellite travel, Rouffet's invention ensures that a fixed point on the Earth's surface likely will remain within a single footprint until it is necessary to switch to another satellite. Because Rouffet's invention does not address handovers caused by the motion of the receiver across the Earth's surface. his arrangement reduces, but does not eliminate, handovers. Figure 3 from the application shows the footprints **12** from six beams aligned in the direction of satellite motion **15**:

FIG. 3



The application contains ten claims that stand or fall as a group. Claim 1 is representative:

> A low orbit satellite communications system for mobile terminals, wherein the communications antenna system of each satellite provides isoflux coverage made up of a plurality of fan beams that are elongate in the travel direction of the satellite.

The examiner initially rejected Rouffet's claims as unpatentable over U.S. Pat. No. 5,199,672 (King) in view of U.S. Pat. No. 4,872,015 (Rosen) and a conference report entitled "A Novel Non-Geostationary Satellite Communications System," *Conference Record*, International Conference on Communications, 1981 (Ruddy). On appeal to the Board, the examiner added an alternative ground for rejection, holding that the claims were obvious over U.S. Pat. No. 5,394,561 (Freeburg) in view of U.S. Pat. No. 5,170,485 (Levine).

On April 16, 1997, the Board issued its decision. Because Rouffet had specified that the claims would stand or fall as a group based on the patentability of claim 1, the Board limited its opinion to that claim. The Board unanimously determined that the examiner had properly rejected claim 1 as obvious over King in view of Rosen and Ruddy. The Board, on a split vote, also affirmed the rejection over Freeburg in view of Levine.

II

To reject claims in an application under section 103, an examiner must show an unrebutted *prima facie* case of obviousness. *See In re Deuel*, 51 F.3d 1552, 1557, 34 USPQ2d 1210, 1214 (Fed. Cir. 1995). In the absence of a proper *prima facie* case of obviousness, an applicant who complies with the other statutory requirements is entitled to a patent. *See In re Oetiker*, 977 F.2d 1443, 1445, 24 USPQ2d 1443, 1444 (Fed. Cir. 1992). On appeal to the Board, an applicant can overcome a rejection by showing insufficient evidence of *prima facie* obviousness or by rebutting the *prima facie* case with evidence of secondary indicia of nonobviousness. *See id.*

While this court reviews the Board's determination in light of the entire record, an applicant may specifically challenge an obviousness rejection by showing that the Board reached an incorrect conclusion of obviousness or that the Board based its obviousness determination on incorrect factual predicates. This court reviews the ultimate determination of obviousness as a question of law. *See In re Lueders*, 111 F.3d 1569, 1571, 42 USPQ2d 1481, 1482 (Fed. Cir. 1997). The factual predicates underlying an obviousness determination include the scope and content of the prior art, the differences between the prior art and the claimed invention, and the level of ordinary skill in the art. *See Monarch Knitting Mach. Corp. v. Sulzer Morat GmbH*, 139 F.3d 877, 881, 45 USPQ2d 1977, 1981 (Fed. Cir. 1998). This court reviews the Board's factual findings for clear error. *See In re Zurko*, 142 F.3d, 1447, 1449, 46 USPQ2d 1691, 1693 (Fed. Cir. 1998) (in banc); *Leuders*, 111 F.3d at 1571-72. " 'A finding is clearly erroneous when, although there is evidence to support it, the reviewing court on the entire evidence is left with the definite and firm conviction that a mistake has been committed.' " *In re Graves*, 69 F.3d 1147, 1151, 36 USPQ2d

1697, 1700 (Fed. Cir. 1995) (quoting *United States v. United States Gypsum Co.*, 333 U.S. 364, 395 [76 USPQ 430] (1948)).

The secondary considerations are also essential components of the obviousness determination. *See In re Emert*, 124 F.3d 1458, 1462, 44 USPQ2d 1149, 1153 (Fed. Cir. 1997) ("Without Emert providing rebuttal evidence, this *prima facie* case of obviousness must stand."). This objective evidence of nonobviousness includes copying, long felt but unsolved need, failure of others, *see Graham v. John Deere Co.*, 383 U.S. 1, 17-18 [148 USPQ 459] (1966), commercial success, *see In re Huang*, 100 F.3d 135, 139-40, 40 USPQ2d 1685, 1689-90 (Fed. Cir. 1996), unexpected results created by the claimed invention, unexpected properties of the claimed invention, *see In re Mayne*, 104 F.3d 1339, 1342, 41 USPQ2d 1451, 1454 (Fed. Cir. 1997); *In re Woodruff*, 919 F.2d 1575, 1578, 16 USPQ2d 1934, 1936-37 (Fed. Cir. 1990), licenses showing industry respect for the invention, *see Arkie Lures, Inc. v. Gene Larew Tackle, Inc.*, 119 F.3d 953, 957, 43 USPQ2d 1294, 1297 (Fed. Cir. 1997); *Pentec, Inc. v. Graphic Controls Corp.*, 776 F.2d 309, 316, 227 USPQ 766, 771 (Fed. Cir. 1985), and skepticism of skilled artisans before the invention, *see In re Dow Chem. Co.*, 837 F.2d 469, 473, 5 USPQ2d 1529, 1532 (Fed. Cir. 1988). The Board must consider all of the applicant's evidence. *See Oetiker*, 977 F.2d at 1445 ("An observation by the Board that the examiner made a *prima facie* case is not improper, as long as the ultimate determination of patentability is made on the entire record."); *In re Piasecki*, 745 F.2d 1468, 1472, 223 USPQ 785, 788 (Fed. Cir. 1984). The court reviews factual conclusions drawn from this evidence for clear error. Whether the evidence presented suffices to rebut the *prima facie* case is part of the ultimate conclusion of obviousness and is therefore a question of law.

When a rejection depends on a combination of prior art references, there must be some teaching, suggestion, or motivation to combine the references. *See In re Geiger*, 815 F.2d 686, 688, 2 USPQ2d 1276, 1278 (Fed. Cir. 1987). Although the suggestion to combine references may flow from the nature of the problem, *see Pro-Mold & Tool Co. v. Great Lakes Plastics, Inc.*, 75 F.3d 1568, 1573, 37 USPQ2d 1626, 1630 (Fed. Cir. 1996), the suggestion more often comes from the teachings of the pertinent references, *see In re Sernaker*, 702 F.2d 989, 994, 217 USPQ 1, 5 (Fed. Cir. 1983), or from the ordinary knowledge of those skilled in the art that certain references are of special importance in a particular field, *see Pro-Mold*, 75 F.3d at 1573 (citing *Ashland Oil, Inc. v. Delta Resins & Refractories, Inc.*, 776 F.2d 281, 297 n.24, 227 USPQ 657, 667 n.24 (Fed. Cir. 1985)). Therefore, "[w]hen determining the patentability of a claimed invention which combines two known elements, 'the question is whether there is something in the prior art as a whole to suggest the desirability, and thus the obviousness, of making the combination.' " *See In re Beattie*, 974 F.2d 1309, 1311-12, 24 USPQ2d 1040, 1042 (Fed. Cir. 1992) (quoting *Lindemann Maschinenfabrik GmbH v. American Hoist & Derrick Co.*, 730 F.2d 1452, 1462, 221 USPQ 481, 488 (Fed. Cir. 1984)).

III

The parties agree that the five references asserted by the examiner are in the same field of endeavor as the invention. The parties also agree that the pertinent level of skill in the art — design of satellite communications systems — is high. On appeal, Rouffet asserts that the examiner and the Board erred by improperly combining references to render the claimed invention obvious.

The Combination of King, Rosen, and Ruddy

The Board first affirmed the rejection of Rouffet's claims over a combination of King, Rosen, and Ruddy. King discloses a system for launching a plurality of satellites into low Earth orbits from a single launch vehicle. Rosen teaches a geostationary satellite that uses a plurality of fan beams with their long axes oriented in an east-west direction to communicate with mobile and fixed terminals on the Earth.

The final, and most important, reference in this combination is Ruddy. Ruddy describes a television broadcast system that uses a series of satellites to retransmit signals sent from a ground station over a wide area. Rather than using a geostationary orbit, Ruddy teaches the use of a series of satellites in Molniya orbits. A satellite in a Molniya orbit always follows the same path through the sky when viewed from a fixed point on the ground. Viewed from the Earth, the orbital path includes a narrow, elliptical apogee loop. In order to transmit to these moving satellites from a ground station, Ruddy uses a fan beam with a long axis aligned with the long axis of the orbit's apogee loop. This alignment places the entire apogee loop within the footprint of the beam and eliminates the need for the ground station's antenna to track the satellite's motion around the apogee loop. Ruddy further teaches orbit param-

eters and spacing of multiple satellites to ensure that a satellite is always in the loop to receive and rebroadcast signals from the Earth station.

King and Rosen together teach the use of a network of satellites in low Earth orbit. Thus, Ruddy becomes the piece of the prior art mosaic that shows, in the reading of the Board, the use of "a plurality of fan beams that are elongate in the travel direction of the satellite." Ruddy, however, is different from the claimed invention in several respects. Specifically, the application claims the projection of multiple elliptical fan-shaped footprints from the satellite to the ground. *See* Claim 1, *supra*, *see also* Application at 6, lines 9-11 ("In addition, in this system, the geometrical shape of the beams 12 is changed: instead of being circular they are now elongate ellipses."). The application's written description further teaches that the invention's fan-shaped satellite beams will minimize handovers. *See id.* at lines 11-16 ("This considerably increases call durations between handovers.").

In contrast, Ruddy teaches that a ground station may use a single fan-shaped beam to transmit to a satellite in a unique Molniya orbit. The ground station transmits a beam into which a series of satellites in Molniya orbits will successively enter. At least two differences are evident: the application teaches projection of multiple beams from a satellite to the Earth, while Ruddy teaches projection of a single beam from the Earth to satellites. Moreover to the extent Ruddy contains a teaching about handovers, its teachings focus on use of the unique Molniya orbit to ensure that a satellite always falls within the beam transmitted by the ground station.

These differences suggest some difficulty in showing a *prima facie* case of obviousness. The Board, however, specifically found that artisans of ordinary skill in this field of art would know to shift the frame of reference from a ground station following a satellite to a satellite transmitting to the ground. According proper deference to the Board's finding of a lofty skill level for ordinary artisans in this field, this court discerns no clear error in the Board's conclusion that these differences would not preclude a finding of obviousness. While Ruddy does not expressly teach alignment of the fan beam with the apparent direction of the satellite's motion, this court perceives no clear error in the Board's determination that Ruddy would suggest such an alignment to one of skill in this art. Therefore, the Board did not err in finding that the combination of King, Rosen, and Ruddy contains all of the elements claimed in Rouffet's application.

[1] However, the Board reversibly erred in determining that one of skill in the art would have been motivated to combine these references in a manner that rendered the claimed invention obvious. Indeed, the Board did not identify any motivation to choose these references for combination. Ruddy does not specifically address handover minimization. To the extent that Ruddy at all addresses handovers due to satellite motion, it addresses this subject through the selection of orbital parameters. Ruddy does not teach the choice of a particular shape and alignment of the beam projected by the satellite. Thus Ruddy addresses the handover problem with an orbit selection, not a beam shape. The Board provides no reasons that one of ordinary skill in this art, seeking to minimize handovers due to satellite motion, would combine Ruddy with Rosen and King in a manner that would render the claimed invention obvious.

Obviousness is determined from the vantage point of a hypothetical person having ordinary skill in the art to which the patent pertains. *See* 35 U.S.C. § 103(a). This legal construct is akin to the "reasonable person" used as a reference in negligence determinations. The legal construct also presumes that all prior art references in the field of the invention are available to this hypothetical skilled artisan. *See In re Carlson*, 983 F.2d 1032, 1038, 25 USPQ2d 1207, 1211 (Fed. Cir. 1993).

As this court has stated, "virtually all [inventions] are combinations of old elements." *Environmental Designs, Ltd. v. Union Oil Co.*, 713 F.2d 693, 698, 218 USPQ 865, 870 (Fed. Cir. 1983); *see also Richdel, Inc. v. Sunspool Corp.*, 714 F.2d 1573, 1579-80, 219 USPQ 8, 12 (Fed. Cir. 1983) ("Most, if not all, inventions are combinations and mostly of old elements."). Therefore an examiner may often find every element of a claimed invention in the prior art. If identification of each claimed element in the prior art were sufficient to negate patentability, very few patents would ever issue. Furthermore, rejecting patents solely by finding prior art corollaries for the claimed elements would permit an examiner to use the claimed invention itself as a blueprint for piecing together elements in the prior art to defeat the patentability of the claimed invention. Such an approach would be "an illogical and inappropriate process by which to determine patentability." *Sensonics, Inc. v. Aerosonic Corp.*, 81 F.3d 1566, 1570, 38 USPQ2d 1551, 1554 (Fed. Cir. 1996).

To prevent the use of hindsight based on the invention to defeat patentability of the invention, this court requires the examiner to

show a motivation to combine the references that create the case of obviousness. In other words, the examiner must show reasons that the skilled artisan, confronted with the same problems as the inventor and with no knowledge of the claimed invention, would select the elements from the cited prior art references for combination in the manner claimed.

[2] This court has identified three possible sources for a motivation to combine references: the nature of the problem to be solved, the teachings of the prior art, and the knowledge of persons of ordinary skill in the art. In this case, the Board relied upon none of these. Rather, just as it relied on the high level of skill in the art to overcome the differences between the claimed invention and the selected elements in the references, it relied upon the high level of skill in the art to provide the necessary motivation. The Board did not, however, explain what specific understanding or technological principle within the knowledge of one of ordinary skill in the art would have suggested the combination. Instead, the Board merely invoked the high level of skill in the field of art. If such a rote invocation could suffice to supply a motivation to combine, the more sophisticated scientific fields would rarely, if ever, experience a patentable technical advance. Instead, in complex scientific fields, the Board could routinely identify the prior art elements in an application, invoke the lofty level of skill, and rest its case for rejection. To counter this potential weakness in the obviousness construct, the suggestion to combine requirement stands as a critical safeguard against hindsight analysis and rote application of the legal test for obviousness.

Because the Board did not explain the specific understanding or principle within the knowledge of a skilled artisan that would motivate one with no knowledge of Rouffet's invention to make the combination, this court infers that the examiner selected these references with the assistance of hindsight. This court forbids the use of hindsight in the selection of references that comprise the case of obviousness. *See In re Gorman*, 933 F.2d 982, 986, 18 USPQ2d 1885, 1888 (Fed. Cir. 1991). Lacking a motivation to combine references, the Board did not show a proper *prima facie* case of obviousness. This court reverses the rejection over the combination of King, Rosen, and Ruddy.

The Combination of Freeburg and Levine

Freeburg teaches a cellular radiotelephone system based on a constellation of low Earth orbit satellites that use conical beams to transmit from the satellite to both fixed and mobile Earth stations. Levine teaches an Earth-based cellular radio system that uses fan beams broadcast from antenna towers. Levine's elliptical footprints are aligned with the road grid. To increase the capacity of traditional ground-based systems through frequency reuse techniques, Levine teaches the use of antennas that broadcast signals with smaller footprints than the prior art system. Thus, Levine actually increases the number of overlap regions between cells and, hence, the number of potential handovers. Figure 1 of the Levine patent illustrates its alignment of beam footprints:

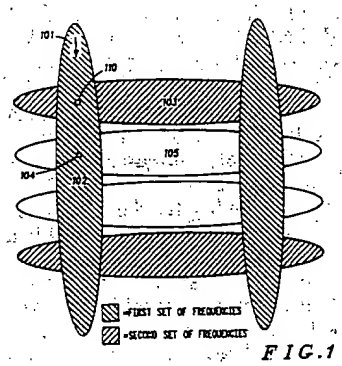


FIG. 1

As a mobile unit (e.g., a driver using a car phone) moves though a succession of overlapping zones, Levine uses selection algorithms to determine which of the cells is aligned with the travel direction of the mobile unit. These algorithms then select this cell for use while continually monitoring intersecting cells in the event that the mobile unit changes direction.

Once again, this court notes significant differences between the teachings of the application and the Levine-Freeburg combination. The critical Levine reference again involves a beam from an Earth station without any reference to the "travel direction of [a] satellite." Moreover, Levine actually multiplies the number of potential handovers and then uses software to sort out the necessary handovers from the unnecessary. However, the Board explains the reasons that one possessing the lofty skills characteristic of this field would know to account for the differences between the claimed invention and the prior art combination. This court discerns no clear error in that reliance on the considerable skills in this field.

[3] This court does, however, discern reversible error in the Board's identification of a motivation to combine Levine and Freeburg. In determining that one of skill in the art would have had motivation to combine Levine and Freeburg, the Board noted that "[t]he level of skill in the art is very high." As noted before, this observation alone cannot supply the required suggestion to combine these references. The Board posits that the high level of skill in the art overcomes the absence of any actual suggestion that one could select part of the teachings of Levine for combination with the satellite system disclosed by Freeburg.

As noted above, the suggestion to combine requirement is a safeguard against the use of hindsight combinations to negate patentability. While the skill level is a component of the inquiry for a suggestion to combine, a lofty level of skill alone does not suffice to supply a motivation to combine. Otherwise a high level of ordinary skill in an art field would almost always preclude patentable inventions. As this court has often noted, invention itself is the process of combining prior art in a nonobvious manner. *See, e.g., Richdel*, 714 F.2d at 1579; *Environmental Designs*, 713 F.2d at 698. Therefore, even when the level of skill in the art is high, the Board must identify specifically the principle, known to one of ordinary skill, that suggests the claimed combination. *Cf. Gechter v. Davidson*, 116 F.3d 1454, 43 USPQ2d 1030 (Fed. Cir. 1997) (explaining that the Board's opinion must describe the basis for its decision). In other words, the Board must explain the reasons one of ordinary skill in the art would have been motivated to select the references and to combine them to render the claimed invention obvious.

The Board's naked invocation of skill in the art to supply a suggestion to combine the references cited in this case is therefore clearly erroneous. Absent any proper motivation to combine part of Levine's teachings with Freeburg's satellite system, the rejection of Rouffet's claim over these references was improper and is reversed.

IV

The Board reversibly erred in determining that there was a motivation to combine either the teachings of King, Rosen, and Ruddy or of Freeburg and Levine in a manner that would render the claimed invention obvious. Because this predicate was missing in each case, the Board did not properly show that these references render the claimed invention obvious. Therefore this court reverses the Board's decision upholding the rejection of Rouffet's claims. In light of this disposition, Rouffet's pending motion to remand the case to the Board for further consideration is denied as moot.

COSTS

Each party shall bear its own costs.

REVERSED.

U.S. Court of Appeals
Federal Circuit

Champagne Louis Roederer S.A. v. Delicato Vineyards

No. 98-1032

Decided July 16, 1998

TRADEMARKS AND UNFAIR TRADE PRACTICES

1. Infringement; conflicts between marks — Likelihood of confusion — Particular marks — Confusion not likely (§335.0304.05)

Infringement; conflicts between marks — Tests generally (§335.06)

Trademark Trial and Appeal Board did not err in dismissing opposition proceeding on ground that applicant's "Crystal Creek" mark for wine and opposer's "Cristal" marks for champagne are dissimilar with respect to appearance, sound, significance, and commercial impression, since board did not err in relying solely on dissimilarity of marks in evaluating likelihood of confusion, since single factor may be dispositive in likelihood of confusion analysis, especially when that factor is dissimilarity of marks, and since no instances of clear error regarding board's findings of fact as to dissimilarities of marks have been demonstrated.

Appeal from the U.S. Patent and Trademark Office, Trademark Trial and Appeal Board.

Proceeding (Opposition No. 80,932) brought by Champagne Louis Roederer S.A. in opposition to application filed by Delicato Vineyards to register mark "Crystal Creek" for wine. From dismissal of opposition proceeding, opposer appeals. Affirmed; Michel, J., concurring in separate opinion.

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- [x] BLACK BORDERS
- [ ] IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- [ ] FADED TEXT OR DRAWING
- [ ] BLURRED OR ILLEGIBLE TEXT OR DRAWING
- [ ] SKEWED/SLANTED IMAGES
- [ ] COLOR OR BLACK AND WHITE PHOTOGRAPHS
- [ ] GRAY SCALE DOCUMENTS
- [ ] LINES OR MARKS ON ORIGINAL DOCUMENT
- [ ] REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- [ ] OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.